PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6:		(11) International Publication Number: WO 98/15648
C12Q 1/68, C07K 14/00	A1	(43) International Publication Date: 16 April 1998 (16.04.98)
(21) International Application Number: PCT/DK (22) International Filing Date: 3 October 1997 (c) (30) Priority Data: 1096/96 4 October 1996 (04.10.96) 1156/96 18 October 1996 (18.10.96) 0512/97 5 May 1997 (05.05.97) (71) Applicant: DAKO A/S [DK/DK]; Produktions DK-2600 Glostrup (DK). (72) Inventors: STENDER, Henrik; Pasanhaven 5, DK-28 tofte (DK). LUND, Kaare; A.D. Jørgensensvej DK-2000 Frederiksberg (DK). MOLLERUP, T. dresen; Lejrevej 14, Allerslev, DK-4320 Lejre (DK Glostrup (DK). (74) Agent: FINK, Kirsten; Dako a/s, Produktionsvej 42, I. Glostrup (DK).	D D D D D D D D D D D D D D D D D D D	CZ, EE, GE, HU, II., IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Buropean patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published With international search report.

(54) Title: NOVEL PROBES FOR THE DETECTION OF MYCOBACTERIA

(57) Abstract

Novel hybridisation assay probes and mixtures of such probes for detecting a target sequence of one or more mycobacteria optionally present in a sample. The probes may suitably be directed to target sequences of mycobacterial rDNA, precursor rRNA, or rRNA, said probes being capable of forming detectable hybrids. The probes are in particular directed to mycobacterial rDNA, to precursor rRNA, or to 23S, 16S or 5S rRNA. The probes are useful for detecting the organisms in test samples such as sputum, laryngeal swabs, gastric lavage, bronchial washings, biopsies, aspirates, expectorates, body fluids (spinal, pleural, pericardial, synovial, blood, pus, bone marrow), urine, tissue sections as well as food samples, soil, air and water samples, and cultures thereof.

BEST AVAILABLE COPY

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
ÄT	Austria	FR	Prance	LU	Luxembourg	SN	Senegal
ΑU	Australia	GA	Gabon	LV	Latvia	8 Z -	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belginm	GN	Guinea	MK	The former Yugoslav	TM	Türkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IR	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL.	Israel	MR	Mauritania	UG	Uganda
BY	Belárus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF.	Central African Republic	JP	Japan	NB	Niger	VN	Viot Nam
ĊG	Congo	KĖ	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	u	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

WO 98/15648 PCT/DK97/00425

NOVEL PROBES FOR THE DETECTION OF MYCOBACTERIA

The present invention relates to novel probes and to mixtures of such probes, in addition to the design, construction and use of such novel probes or mixtures thereof for detecting a target sequence of one or more mycobacteria, which probes are capable of detecting such organism(s) optionally present in a test sample, e.g. sputum, laryngeal swabs, gastric lavage, bronchial washings, biopsies, aspirates, expectorates, body fluids (spinal, pleural, pericardial, synovial, blood, pus, bone marrow), urine, tissue sections as well as food samples, soil, air and water samples and cultures thereof. The invention relates in particular to novel probes and mixtures thereof for detecting the presence of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) and for detecting the presence of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT). The invention further relates to diagnostic kits comprising one or more of such probes. The probes of the present invention are surprisingly able to penetrate the cell wall of the mycobacteria, thus making possible the development of fast an easy-performed in situ protocols.

BACKGROUND OF THE INVENTION

5

10

15

20

25

30

35

Tuberculosis is a very life-threatening and highly epidemic disease which is caused by infection with Mycobacterium tuberculosis. Tuberculosis is presently the predominant infectious cause of morbidity and mortality world-wide, and is estimated to kill about three million people annually. WHO estimates that the annual number of new cases of tuberculosis will increase from 7.5 million in 1990 to 10.2 million in 2000, an escalation that will result in approximately 90 million new cases during this decade. It is furthermore estimated that 30 million people will die from tuberculosis during the 1990s, which equals one quarter of preventable deaths among adults.

The prevalence of tuberculosis has been very high in the poorer parts of the world such as Asia, Africa and South-America, but in recent years an increase has also been observed in industrialised countries. This appears to be due to an interaction of various factors including i.a. patterns of migration, poorly organised tuberculosis programmes and nutrition problems. Furthermore, a serious threat will arise from the emergence of new strains that are drug resistant or worse, multi-drug resistant.

Mycobacteria are often divided into tuberculous mycobacteria, i.e. mycobacteria of the Mycobacterium tuberculosis Complex (MTC), and non-tuberculous mycobacteria, i.e. mycobacteria other than those of the Mycobacterium tuberculosis Complex (MOTT). The MTC

group comprises apart from M. tuberculosis, M. bovis, M. africanum and M. microti.

Mycobacteria of the MOTT group are not normally pathogenic to healthy individuals but may cause disease in immunocompromised individuals, e.g. individuals infected with HIV. Clinical relevant mycobacteria of the MOTT group are in particular M. avium, M. intracellulare, M. kansasii and M. gordonae, but also M. scrofulaceum, M. xenopi and M. fortuitum.

M. avium and M. intracellulare together with M. paratuberculosis and M. lepraemurium constitute the Mycobacterium avium Complex (MAC). Extended with M. scrofulaceum, the group is named Mycobacterium avium -intracellulare -scrofulaceum Complex (MAIS).

10

15

20

30

35

It is well-known that treatment of mycobacterial infections with antibiotics may lead to the emergence of drug resistant strains. Many antibiotic drugs excert their effects by interfering with protein synthesis or with transcription. Studies of the molecular mechanisms underlying certain antibiotic resistance phenotypes in clinical mycobacterium isolates have revealed mutations in rRNA genes. The development of resistance because of mutation(s) located in the rRNA gene is likely to occur since slow-growing mycobacteria have only a single rRNA operon. All mycobacteria populations comprise a minority of drug resistant mutants that have arisen by spontaneous mutation. These mutated mycobacteria do normally not survive particularly well, but, when single-drug therapy is offered as treatment, the drug susceptible bacteria are killed, and only the resistant mutants will survive and multiply, and, thus at some point, constitute the majority of the mycobacterial population. The selection of drug resistant bacteria due to inadequate drug therapy leads to a state of so-called "acquired drugresistance". In contrast, "primary drug-resistance" is used to characterise a situation where drug-resistant mycobacteria can be isolated from a patient who has never been treated for mycobacterial infection, and has become infected with drug-resistant mycobacteria from an individual suffering from infection with an acquired drug resistant bacterium.

Today, drug-resistance is determined primarily phenotypically by culturing clinical samples, in which presence of mycobacteria have been demonstrated, in the presence of the individual drugs. This is unfortunately a very slow and time-consuming procedure as the result of the drug-resistance studies depends on the growth rate of the mycobacteria, which are well-known to be slow. Thus, the result is not available until after several weeks.

Although the incidence of drug-resistance is, at least not yet, very common, it is nevertheless very important that resistant strains are identified and eradicated. Therefore, it is of major importance to find a reliable and rapidly performed method of diagnosing drug-resistance.

Presently, the detection of mycobacteria by microscopy is the most prevalent method for

diagnosis. The sample (e.g. an expectorate) is stained for the presence of acid-fast bacilli using e.g. Ziehl-Neelsen staining. However, staining for acid-fast bacilli does not provide the necessary information about the type of infection, only whether acid fast bacilli are present in the sample, and this is in itself not sufficient information for establishing a diagnosis. Samples positive for acid fast bacilli, may subsequently be cultured in order to be able to perform species identification.

5

10

15

20

25

30

35

Since Ziehl-Neelsen staining cannot be used to determine whether the infection is caused by mycobacteria of the MTC group or mycobacteria other than mycobacteria of the MTC group, a positive staining frequently leads to very costly isolation of all the patients with suspected M. tuberculosis infection as well as treatment with medicaments to which the patient may not even respond.

Since the sensitivity of acid fast staining is only approximately 10⁴-10⁵ per ml smear negative samples should also be cultured as culture-based tests are sensitive, and as it may be possible to detect 10-100 organisms per sample, but the result is not available before up to 8 weeks of culturing. Likewise, information about drug susceptibility is not available until after 1-3 weeks of further testing.

Different solid or liquid media (Loewenstein Jensen slants and Dubos broth) have traditionally been used for culturing of mycobacteria-containing samples. Newer media include ESP Myco Culture System (Difco), MB/BacT (Organon Teknika), BacTec (Becton Dickinson) and MGIT (Becton Dickinson). These test media are based on colourmetric or fluorometric detection of carbon dioxide or oxygen produced by mycobacterial metabolism, and adapted to automated systems for large scale testing.

Species identification is presently carried out following culturing using traditional biochemical methods or probe hybridisation assays (e.g. AccuProbe by Gen-Probe Inc., USA). There is, therefore, an increasing need for means allowing a more rapid distinction between mycobacteria of the MTC group and mycobacteria other than those of the MTC group, and for further species identification of those especially mycobacteria other than those of the MTC group.

A number of new attempts to replace the culture-based methods relies on molecular amplification technology. Several methods have emerged, among them the polymerase chain reaction (PCR), the ligase chain reaction and transcription mediated amplification. The basic principle of amplification methods is that a specific nucleic acid sequence of the mycobacteria is amplified to increase the copy number of the specific sequence to a level where the

PCT/DK97/00425

amplicon may be detectable. In principle, the methods offers the possibility of detecting only one target sequence, thus, in principle, making detection of mycobacteria present at low levels possible. However, it has become clear that the target amplification methods cannot replace culture-based methods as only samples which are positive by staining for acid fast bacilli (AFB) give a satisfactory sensitivity. Furthermore, specific problems exist for each method. The PCR method may give false negative results due to the presence of inhibitors such as haemoglobin. Another problem arises from cross-contamination of negative specimens and/or reagents with amplified nucleic acid present in the laboratory environment leading to false positive results. A disadvantage is that costly reagents are needed for performing these tests. Furthermore, specialised instrumentation is required, making these tests mainly useful in large specialised laboratories, and generally not applicable in smaller clinical laboratories.

Nucleic acid probes for detecting rRNA of mycobacteria have been described in for example US 5 547 842, EP-A 0 572 120 and US 5 422 242.

15

10

Considering the perspective and impact the disease has, the development of rapid and preferably easy-performed and further economic feasible diagnostic detection tests are of utmost importance and would be a very valuable tool in the fight against the spread of tuberculosis.

20

Peptide nucleic acids are pseudo-peptides with DNA-binding capability. The compounds were first reported in the early nineties in connection with a series of attempts to design nucleotide analogues capable of hybridising, in a sequence-specific fashion, to DNA and RNA, cf. WO 92/20702.

25

30

35

Hybridisation of peptide nucleic acid probes to DNA and to RNA has been shown to obey the Watson-Crick base pairing rules, and peptide nucleic acid probes have been found to hybridise to a DNA or a RNA target with higher affinity and specificity than the nucleic acid counterparts. These properties are ascribed to the uncharged, as opposed to the charged, structure of the peptide nucleic acid and DNA or RNA backbones, respectively, and to the high conformational flexibility of the peptide nucleic acid molecules. These features - together with the documented stability of peptide nucleic acid towards a variety of naturally occurring nucleases and proteases that usually degrade DNA, RNA or proteins - invite for use of peptide nucleic acid probes as antisense therapeutic agents and opens potentially important applications in diagnostics.

SUMMARY OF THE INVENTION

10

15

20

25

30

35

The present invention relates to novel peptide nucleic acid probes and to mixtures of such probes for detecting a target sequence of one or more mycobacteria optionally present in a sample. In accordance with claim 1, the probes are directed to target sequences of mycobacterial rRNA, genomic sequences corresponding to said rRNA (rDNA) and precursor rRNA. rRNA is present in a high number of copies in each cell, and is hence a well suited target. The probes are, as defined in claim 2, suitably directed to target sequences of mycobacterial rDNA, precursor rRNA, or 23S, 16S or 5S rRNA.

Thus, in a first aspect, the invention features a hybridisation assay probe and a mixture of such probes for detecting a target sequence of one or more mycobacteria in accordance with claim 1 and 2. Under appropriate stringency conditions, Such probes should not to any significant degree cross-react with ribosomal nucleic acid from other not relevant organisms, present in the test sample, in particular other mycobacteria. Cross-reactivity to organisms that are unlikely to be present in the sample may not be of importance. In in situ assays implying examination by microscopy, it is further possible to distinguish mycobacteria from other bacteria based on the morphology of these bacilli.

The invention also relates to peptide nucleic acid probes in accordance with claim 3 for obtaining a target sequence and in accordance with claim 4 for obtaining a probe.

In another aspect, the invention relates to novel peptide nucleic acid probes for detecting a target sequence of one or more mycobacteria of the MTC group, and one or more mycobacteria other than mycobacteria of the MTC group, which probes comprise from 6 to 30 polymerised peptide nucleic acid moieties (claim 5). Suitable probes of formula (I) are claimed in claim 6.

Claims 7 to 10 and 15 to 24 relate to probes or mixtures of such probes for detecting a target sequence of one or more mycobacteria of the MTC group. Claims 11 to 13 and 15 to 24 relate to probes or mixtures of such probes for detecting a target sequence of one or more mycobacteria other than mycobacteria of the MTC group (MOTT group). Claim 14 relates specifically to probes for detecting drug resistant mycobacteria. Claims 25 to 27 relate to the use of such probes or mixtures thereof.

In accordance with claims 28 to 34, the present invention also relates to a method for detecting the presence of mycobacteria.

In yet another aspect, the present invention relates to a kit (claim 35 and 36) comprising at least one peptide nucleic acid probe as defined in claims 1 to 24.

Mycobacteria are characterised by a complex cell wall which contains myolic acids, complex waxes and unique glycolipids. It is generally recognised by those skilled in the art that this wall provides mycobacteria with extreme resistance to chemical and physical stress as compared to other bacteria, and, accordingly, makes them very difficult to penetrate and lyse. The low permeability of the cell wall is considered to be the main reason for the fact that only very few drugs are effective in the treatment of tuberculosis and other mycobacterial infections. As described in US 5 582 985, the wall appears further to prevent penetration by nucleic acid probes. Even with short probes (shorter than 30 nucleic acids), specific staining is low or often non-existent. Protocols that allow DNA probes to be used for in situ hybridisation to mycobacterial species are described in US 5 582 985. However, these protocols require dewaxing of the mycobacterial cell wall with xylene and further enzymatic treatment prior to the hybridisation step in order to make the mycobacterial cell wall permeable to the DNA probes.

15

20

25

10

The problems set forth above have surprisingly been completely solved by the use of peptide nucleic acid probes. It has, surprisingly, been found that the peptide nucleic acid probes are able to penetrate the cell wall of the mycobacteria, and furthermore that this is taking place rapidly. The person skilled in the art would arrive at the conviction that it would be necessary to heavily treat the mycobacteria before hybridisation is carried out. Thus, based on the available prior art, there is a strong prejudice against carrying out hybridisation without prior destruction of the mycobacterial cell wall. The inventors have shown that this is indeed and unexpectedly possible. It has been demonstrated that the probes of the present invention are able to hybridise to mycobacterial precursor rRNA and rRNA without harsh treatment of the mycobacterial cells, thus avoiding a risk of interfering with the morphology of the cells. Using the present probes, specific and easy detection and, subsequently, diagnosis of tuberculosis and other mycobacterial infections are thus possible.

BRIEF DESCRIPTION OF THE FIGURES

30

35

Alignments of rDNA sequences of M. tuberculosis (as a representative of the MTC group) and important closely related species thereto, including M. avium (as a representative of the mycobacteria other than those of the MTC group) and important closely related species thereto for the 23S, 16S and/or 5S rRNA genes have been made (Figures 1A-1J, 2A-2D, 3, 4A-4L and 5A-B). The alignment for M. bovis and M. intracellulare are partly based on public available sequences and partly on sequences obtained by sequencing performed at DAKO A/S.

10

15

Alignment for the MTC group (23S rDNA)

Figures 1A-1J show alignments of 23S rDNA sequences of M. tuberculosis (GenBank entry GB:MTCY130, accession number Z73902), M. avium (GenBank entry GB:MA23SRNA, accession number X74494), M. paratuberculosis (GenBank entry GB:MPARRNA, accession number X74495), M. phlei (GenBank entry GB:MP23SRNA, accession number X74493), M. leprae (GenBank entry GB:ML5S23S, accession number X56657), M. gastri (GenBank entry GB:MG23SRRNA, accession number Z17211), M. kansasii (GenBank entry GB:MK23SRRNA, accession number Z17212), and M. smegmatis (GB:MS16S23S5, accession number Y08453). Preferred peptide nucleic acid probes should enclose at least one nucleobase complementary to a nucleobase of M. tuberculosis 23S rRNA within positions 149-158, 220-221, 328-361, 453-455, 490-501, 637-660, 706-712, 762-789, 989, 1068-1072, 1148, 1311-1329, 1361-1364, 1418, 1563-1570, 1627-1638, 1675-1677, 1718, 1734-1740, 1967-1976, 2403-2420, 2457-2488, 2952-2956, 2966-2969, 3000-3003, and 3097-3106 of the alignment (indicated by heavy frames). Differences between the sequences of M. avium, M. phlei, M. leprae, M. paratuberculosis, M. gastri and M. kansasii and that of M. tuberculosis in the alignment are indicated by light frames.

Alignment for the MTC group (16S rDNA)

Figures 2A-2D show alignments of 16S rDNA sequences of M. tuberculosis (GenBank entry GB:MTU16SRN, accession number X52917), M. bovis (GenBank entry GB:MSGTGDA, 20 accession number M20940), M. avium (GenBank entry GB:MSGRRDA, accession number M61673), M. intracellulare (GenBank entry GB:MIN16SRN, accession number X52927), M. paratuberculosis (GenBank entry GB:MSGRRDH, accession number M61680), M. scrofulaceum (GenBank entry GB:MSC16SRN, accession number X52924), M. leprae 25 (GenBank entry GB:MLEP16S1, accession number X55587), M. kansasii (GenBank entry GB:MKRRN16, accession number X15916), M. gastri (GenBank entry GB:MGA16SRN, accession number X52919), M. gordonae (GenBank entry GB:MSGRR16SI, accession number M29563) and M. marinum (GenBank entry GB:MMA16SRN, accession number X52920). Preferred peptide nucleic acid probes should enclose at least one nucleobase 30 complementary to a nucleobase of M. tuberculosis 16S rRNA within positions 76-79, 98-101, 135-136, 194-201, 222-229, 242, 474, 1136-1145, 1271-1272, 1287-1292, 1313, and 1334 of the alignment (indicated by heavy frames). Differences between the sequences of M. bovis, M. avium, M. intracellulare, M. paratuberculosis, M. scrofulaceum, M. leprae, M. kansasii, M. gastri, M. gordonae and M. marinum, and that of M. tuberculosis in the alignment are indicated 35 by light frames.

Alignment for the MTC group (5S rDNA)

Figure 3 shows alignments of 5S rDNA sequences of M. tuberculosis (GenBank entry

GB:MTDNA16S, accession number x75601), M. bovis (GenBank entry GB:MBRRN5S, accession number X05526), M. phlei (GenBank entry GB:MP5SRRNA, accession number X55259), M. leprae (GenBank entry GB:ML5S23S, accession number X56657), and M. smeamatis (GenBank entry GB:MS16S23S5, accession number Y08453). Preferred peptide nucleic acid probes should enclose at least one nucleobase complementary to a nucleobase of M. tuberculosis 5S rRNA within positions 86-90 of the alignment (indicated by heavy frame). Differences between the sequences of M. bovis, M. phlei, M. leprae, M. smegmatis and M. luteus and that of M. tuberculosis in the alignment are indicated by light frames.

5

- Alignment for Mycobacteria other than those of the MTC group (23S rDNA) 10 Figures 4A-4L show alignments of 23S rDNA sequences of M. avium (GenBank entry GB:MA23SRNA, accession number X74494), M. paratuberculosis (GenBank entry GB:MPARRNA, accession number X74495), M. tuberculosis (GenBank entry GB:MTCY130, accession number Z73902), M. phlei (GenBank entry GB:MP23SRNA, accession number 15 X74493), M. leprae (GenBank entry GB:ML5S23S, accession number X56657), M. gastri (GenBank entry GB:MG23SRRNA, accession number Z17211), M. kansasii (GenBank entry GB:MK23SRRNA, accession number Z17212), and M. smegmatis (GB:MS16S23S5, accession number Y08453). Preferred peptide nucleic acid probes should enclose at least one nucleobase complementary to a nucleobase of M. avium 23S rRNA within positions 99-101, 20 183, 261-271, 281-284, 290-293, 327-335, 343-357, 400-405, 453-462, 587-599, 637-660. 704-712, 763-789, 1060-1074, 1177-1185, 1259-1265, 1311-1327, 1345-1348, 1361-1364, 1556-1570, 1608-1613, 1626-1638, 1651-1659, 1675-1677, 1734-1741, 1847-1853, 1967-1976, 2006-2010, 2025-2027, 2131-2232, 2252-2255, 2396-2405, 2416-2420, 2474-2478 2687, 2719, 2809, 3062-3068, and 3097-3106 of the alignment (indicated by heavy frames). Differences between the sequences of M. paratuberculosis, M. tuberculosis, M. phlei, M. leprae, M. gastri, M. kansasii, and M. smegmatis and that of M. avium in the alignment are indicated by light frames.
- Alignment for Mycobacteria other than those of the MTC group (16S rDNA) 30 Figures 5A-5B show alignments of 16S rDNA sequences of M. avium (GenBank entry GB:MSGRRDA, accession number M61673), M. intracellulare (GenBank entry GB:MIN16SRN, accession number X52927),M. paratuberculosis (GenBank entry GB:MSGRRDH, accession number M61680), M. scrofulaceum (GenBank entry GB: MSC16SRN, accession number X52924), M. tuberculosis (GenBank entry GB:MTU16SRN, accession number X52917), M. bovis (GenBank entry GB:MSGTGDA, accession number M20940), M. leprae (GenBank entry GB:MLEP16S1, accession number X55587), M. kansasii (GenBank entry GB:MKRRN16, accession number X15916), and M. gastri (GenBank entry GB:MGA16SRN, accession number X52919), M. gordonae (GenBank entry GB:MSGRR16SI,

accession number M29563), and M. marinum (GenBank entry GB:MMA16SRN, accession number X52920). Preferred peptide nucleic acid probes should enclose at least one nucleobase complementary to a nucleobase of M. avium 16S rRNA within positions 135-136, 472-475, 1136-1144, 1287-1292, 1313, and 1334 of the alignment (indicated by heavy frames). Differences between the sequences of M. intracellulare, M. paratuberculosis, M. scrofulaceum, M. tuberculosis, M. bovis, M. leprae, M. kansasii, and M. gastri and that of M. avium in the alignment are indicated by light frames.

Drug-resistance

5

20

25

30

- Figure 6 shows a partial M. avium 23S rDNA sequence including positions 2550 to 2589 of GenBank entry X74494. Bases in positions where deviations from the wild-type sequence have been correlated with macrolide-resistance are framed. Positions 2568 and 2569 in the figure correspond to positions 2058 and 2059, respectively, of E. coli 23S rRNA.
- Figure 7 shows a partial M. tuberculosis 16S rDNA sequence including positions 441 to 491 and 843 to 883 of GenBank entry X52917. Bases in positions where deviations from the wild-type sequence have been correlated with resistance to streptomycin are framed. Positions 452, 473, 474, 477, 865, and 866 in the figure correspond to positions 501, 522, 523, 526, 912, and 913, respectively, of E.coli 16S rRNA.

SPECIFIC DESCRIPTION

characteristics.

The present invention provides novel probes for use in rapid and specific, sensitive hybridisation based assays for detecting a target sequence of one or more mycobacteria, which target sequence is located in the mycobacterial rDNA, precursor rRNA, or in the 23S, 16S or 5S rRNA. The probes to be used in accordance with the present invention are peptide nucleic acid probes. Peptide nucleic acids are non-naturally occurring polyamides or polythioamides which can bind to nucleic acids (DNA and RNA). Such compounds are described in e.g. WO 92/20702.

We have identified suitable variable regions of the target nucleic acid by comparative analysis of generally available rDNA sequences and sequences obtained by sequencing as described above. Computers and computer programs, which have been used for the purposes disclosed herein, are commercially available. From such alignments, possibly suitable probes can be identified. The alignments are thus a useful guideline for designing probes with desired

When designing the probes, due regard should be taken to the assay conditions under which

5

10

15

20

25

30

35

the probes are to be used. Stringency is chosen so as to maximise the difference in stability between the hybrid formed with the target nucleic acid and that formed with the non-target nucleic acid. It will typically be necessary to choose high stringency conditions for probes where the specificity depends on only one mismatch to non-target sequences. The more mismatches to non-target sequences, the less demand for high stringency conditions.

PCT/DK97/00425

Furthermore, probes should be designed so as to minimise the stability of probe-non-target nucleic acid hybrids. This may be accomplished by minimising the degree of complementarity to non-target nucleic acid, i.e. by designing the probe to span as many destabilising mismatches as possible, and/or to include as many additions/deletions relative to the target sequence as possible. Whether a probe is useful for detecting a particular mycobacterial species depends to some degree on the difference between the thermal stability of probetarget hybrids and probe:non-target hybrids. For rRNA targets, however, the secondary structure of the region of the rRNA molecule in which the target sequence is located may also be of importance. The secondary structure of a probe should also be taken into consideration. Probes should be designed so as to minimise their proclivity to form hairpins, self-dimers, and pair-dimers if a mixture of two or more probes is used.

Mismatching bases in hybrids formed between peptide nucleic acid probes and nucleic acids result in a higher thermal instability than mismatching bases in nucleic acid duplexes of the same sequences. Thus, the peptide nucleic acid probes exhibit a greater specificity for a given target nucleic acid sequence than a traditional nucleic acid probe, which is seen as a greater difference in T_m values for probe-target hybrids and probe-non-target hybrids. The sensitivity and specificity of a peptide nucleic acid probe will also depend on the hybridisation conditions used.

The primary concern regarding the length of the peptide nucleic acid probes is the warranted specificity, i.e. which length provides sufficient specificity for a particular application. The optimal length of a peptide nucleic acid probe comprising a particular site with differences in base composition, e.g. among selected regions of mycobacterial rRNA, is a compromise between the general pattern that longer probes ensure specificity and shorter probes ensure that the destabilising differences in base composition constitute a greater portion of the probe. Also, due regard must be paid to the conditions under which the probes are to be used.

Peptide nucleic acid sequences are written from the N-terminal end of the sequence towards the C-terminal end. A free (unsubstituted) N-terminal end or an N-terminal end terminating with an amino acid is indicated as H, and a free C-terminal end is indicated as NH₂ (a carboxamide group). Peptide nucleic acids are capable of hybridising to nucleic acid

sequences in two orientations, namely in antiparallel orientation and in parallel orientation. The peptide nucleic acid is said to hybridise in the antiparallel orientation when the N-terminal end of the peptide nucleic acid is facing the 3' end of the nucleic acid sequence, and to hybridise in the parallel orientation when the C-terminal end of the peptide nucleic acid is facing the 5' end of the nucleic acid sequence. In most applications, hybridisation in the antiparallel orientation is preferred as the hybridisation in the parallel orientation takes place rather slowly and as the formed duplexes are not as stable as the duplexes having antiparallel strands. Triplex formation with a stoichiometry of two peptide nucleic acid strands and one nucleic acid strand may occur if the peptide nucleic acid has a high pyrimidine content. Such triplexes are very stable, and probes capable of forming triplexes may thus be suitable for certain applications.

Mainly because the peptide nucleic acid strand is uncharged, a peptide nucleic acid-nucleic acid-duplex will have a higher T_m than the corresponding nucleic acid-nucleic acid-duplex. Typically there will be an increase in T_m of about 1 °C per base pair at 100 mM NaCl depending on the sequence (Egholm et al. (1993), Nature, 365, 566-568).

10

15

20

25

30

35

In contrast to DNA-DNA-duplex formation, no salt is necessary to facilitate and stabilise the formation of a peptide nucleic acid-DNA or a peptide nucleic acid-RNA duplex. The T_m of the peptide nucleic acid-DNA-duplex changes only little with increasing ionic strength. Typically for a 15-mer, the T_m will drop only 5 °C when the salt concentration is raised from 10 mM NaCl to 1 M NaCl. At low ionic strength (e.g. 10 mM phosphate buffer with no salt added), hybridisation of a peptide nucleic acid to a target sequence is possible under conditions where no stable DNA-DNA-duplex formation occurs. Furthermore, target sites that normally are inaccessible can be made more readily accessible for hybridisation with peptide nucleic acid probes at low salt concentration as the secondary and tertiary structure of nucleic acids are destabilised under such conditions. Using peptide nucleic acid probes, a separate destabilising step or use of destabilising probes may not be necessary to perform.

rRNA is essential for proper function of the ribosomes and thus the synthesis of proteins. The genes encoding the rRNAs are in eubacteria located in an operon in which the small subunit RNA gene, the 16S rRNA gene, is located nearest the 5' end of the operon, the gene for the large subunit RNA, the 23S rRNA gene, is located distal to the 16S rRNA gene and the 5S rRNA gene is located nearest the 3' end of the operon. The three genes are separated by spacer regions in which tRNA genes may be found, however, there are none in M. tuberculosis. The primary transcript of the eubacterial rRNA operon is cleaved by RNaselll. This cleavage results in separation of the 16S, the 23S and the 5S rRNA into precursor rRNA molecules (pre-rRNA molecules) which besides the rRNA species also contain leader and tail sequences. The primary RNase III cleavage is normally a rapid process, whereas the

WO 98/15648 PCT/DK97/00425

subsequent maturation is substantially slower. Precursor rRNA is typically more abundant than even strongly expressed mRNA species. Thus, for certain applications, precursor rRNA may be an attractive diagnostic target. In order to specifically detect precursor rRNA, a target probe should be directed against sequences comprising at least part of the leader or tail sequences. A target probe may further be directed against sequences of which both part of the leader/tail and mature rRNA sequences are constituents.

Usually, patients having contracted a mycobacterial infection are treated with medicaments until no mycobacteria can be found in the sputum. Except for culturing, the presently available methods do not allow for clear distinguishing between living and dead mycobacteria. This means that a patient may often be treated with medicaments for a longer period of time than actually necessary. A way of determining the progress of treatment would be a very valuable tool in the fight of tuberculosis and other mycobacterial diseases.

As transcription and maturation of rRNA is a measure of viability, detection of precursor rRNA is a suitable and direct measure of viability of the bacteria. Furthermore, precursor rRNA may be used for identification of antibiotic drugs which reduce or inhibit rRNA transcription. One such example is rifampicin. A transcriptional inhibitor will in susceptible bacteria eliminate new synthesis of rRNA and thus the pool of precursor rRNA will be depleted. However, in resistant cells, primary transcripts as well as precursor rRNAs will continue to be produced.

Although it is preferred to use peptide nucleic acid probes targeting specific sequences of rRNA, it will readily be understood that peptide nucleic acid probes complementary to rRNA targeting probes will be useful for the detection of the genes coding for said sequence specific rRNA (rDNA), and peptide nucleic acid probes for the detecting rDNA is hence contemplated by the present invention. Although it is preferred to choose the sequence of the probe so as to enable the probe to hybridise to its target sequence in antiparallel orientation, it is to be understood that probes capable of hybridising in parallel orientation can be constructed from the same information. The present invention is intended to cover both types of probes.

30

25

5

10

15

20

In the broadest aspect, the present invention relates to peptide nucleic acid probes for detecting a target sequence of one or more mycobacteria optionally present in a test sample, said probe being capable of hybridising to a target sequence of mycobacterial rDNA, precursor rRNA or rRNA (claim 1).

35

The probes of the invention may suitably be directed to rDNA, precursor rRNA, or to 23S, 16S or 5S rRNA.

In accordance with claim 3, the target sequences, to which the peptide nucleic acid probe(s) are capable of hybridising to, are obtainable by

13

- (a) comparing the nucleobase sequences of said mycobacterial rRNA or rDNA of one or more mycobacteria to be detected with the corresponding nucleobase sequence of organism(s), in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished,
 - (b) selecting a target sequence of said rRNA or rDNA which includes at least one nucleobase differing from the corresponding nucleobase of the organism(s), in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished, and (c) determining the capability of said probe to hybridise to the selected target sequence to form detectable hybrids.

Peptide nucleic acid probes are, in accordance with claim 4, obtainable by

15

30

10

- (a) comparing the nucleobase sequences of said mycobacterial rRNA or rDNA of one or more mycobacteria to be detected with the corresponding nucleobase sequence of organism(s), in particular other mycobacteria, in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished,
- (b) selecting a target sequence of said rRNA or rDNA which includes at least one nucleobase differing from the corresponding nucleobase of the organism(s), in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished,
 (c) synthesising said probe, and
- (4) determining the capability of said probe to hybridise to the selected target sequence toform detectable hybrids.

The probes are in particular suitable for detecting a target sequence of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) or for detecting a target sequence of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT) optionally present in a sample, which probe comprises from 6 to 30 polymerised peptide nucleic acid moieties, said probe being capable of hybridising to a target sequence of mycobacterial rDNA, precursor rRNA or 23S, 16S or 5S rRNA forming detectable hybrids (claim 5). In accordance with claim 6, such probes may comprise peptide nucleic acid moieties of formula (I)

WO 98/15648 PCT/DK97/00425

wherein each X and Y independently designate O or S, each Z independently designates O, S, NR^1 , or $C(R^1)_2$, wherein each R^1 independently designate H, C_{1-6} alkyl, C_{1-6} alkyl, C_{1-6} alkynyl,

each R², R³ and R⁴ designate independently H, the side chain of a naturally occurring amino acid, the side chain of a non-naturally occurring amino acid, C₁₋₄ alkyl, C₁₋₄ alkenyl or C₁₋₄ alkynyl, or a functional group, each Q independently designates a naturally occurring nucleobase, a non-naturally occurring nucleobase, an intercalator, a nucleobase-binding group, a label or H,

15 and with the proviso indicated in claim 6.

5

10

20

25

30

35

The probes may suitably be used for detecting a species specific mycobacterial target sequence, or target sequences of a group of mycobacteria like MTC, MOTT, MAC or MAIS. The probes may further be designed so as to be capable of hybridising to one or more drug resistant mycobacteria, or, alternatively, to the wild-type corresponding thereto. In the design of the probes, sequences between different mycobacteria (one or more) may be taken into account as may sequences from other related or non-related organisms (one or more).

As mentioned above, drug-resistance is an increasing threat to the fight of mycobacterial infection. Monotherapy with macrolides such as clarithromycin and azithromycin often leads to clinically significant drug-resistance. Clarithromycin and azithromycin are important drugs in the treatment of infections by especially M. avium. Comparison between 23S rRNA sequences from isolates of M. avium and M. intracellulare with acquired resistance to clarithromycin and azithromycin and 23S rRNA sequences from non-resistant strains has revealed that the majority of resistant strains have single-point mutations in the 23S rRNA in positions corresponding to 2058 and 2059 in E. coli 23S rRNA. In the M. avium 23S rRNA sequence (GenBank accession number X74494), the corresponding bases are in position 2568 and 2569, respectively, as shown in Figure 6. Most susceptible strains have an A residue in one of these positions whereas the resistant strains have a C, G or T in position 2058 (E. coli numbering corresponding to 2568 in M. avium with GenBank accession number X74494), or G or C in position 2059 (E. coli numbering corresponding to 2569 in M. avium with GenBank accession number X74494).

Single-point mutations in rRNA apparently also account to some degree for streptomycin resistance. Streptomycin, the first successful antibiotic drug against tuberculosis, is an aminocyclitol glycoside that perturbs protein synthesis at the ribosomal level. The genetic basis for streptomycin resistance has not yet been completely solved. However, some streptomycin resistant strains of M. tuberculosis have single-point mutations in 16S rRNA. These mutations are located in positions corresponding to bases 501, 522, 523, 526, 912 and 913 in E. coli 16S rRNA which correspond to bases with numbers 452, 473, 474, 477, 865 and 866, respectively, of M. tuberculosis 16S rRNA (GenBank accession number X52917) as shown in Figure 7. Most of these mutated nucleotides are involved in structural interactions within the 530 loop of 16S rRNA which is one of the most conserved regions in the entire 16S rRNA gene.

5

10

15

20

25

30

35

Mutations in an 81 bp region of the gene (rpoB) encoding the beta subunit of RNA polymerase are the cause of 96% of the cases of rifampicin resistance in M. tuberculosis and M. leprae. rRNA precursor molecules have terminal domains (tails) which are removed during late steps in precursor rRNA processing to yield the mature rRNA molecules. Transcriptional inhibitors such as rifampicin can deplete precursor rRNA in sensitive cells by inhibiting de novo precursor rRNA synthesis while allowing maturation to proceed. Thus, precursor rRNA is depleted in susceptible mycobacterium cells while it remains produced in resistant mycobacterium cells when the cells are exposed to rifampicin during culturing.

Peptide nucleic acid probes for detecting a target sequence of one or more mycobacteria of the Mycobacterium tuberculosis Complex are defined in claims 7 to 10. Peptide nucleic acid probes for detecting a target sequence of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex are defined in claims 11 to 13. Peptide nucleic acid probes for detecting a target sequence of one or more drug resistant mycobacteria of the Mycobacterium tuberculosis complex or of one or more drug resistant mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex are defined in claim 14.

In the present context and the claims, the term "naturally occurring nucleobases" includes the four main DNA bases (i.e. thymine (T), cytosine (C), adenine (A) and guanine (G)) as well as other naturally occurring nucleobases (e.g. uracil (U) and hypoxanthine).

The term "non-naturally occurring nucleobases" comprises i.a. modified naturally occurring nucleobases. Such non-naturally occurring nucleobases may be modified by substitution by e.g. one or more C₁₋₈ alkyl, C₁₋₈ alkenyl or C₁₋₈ alkynyl groups or labels. Examples of non-naturally occurring nucleobases are purine, 2,6-diamino purine, 5-propynylcytosine (C propynyl), isocytosine (iso-C), 5-methyl-isocytosine (iso-C) (see e.g. Tetrahedron Letters Vol

35

36, No 12, 2033-2036 (1995) or Tetrahedron Letters Vol 36, No 21, 3601-3604 (1995)), 7-deazaadenine, 7-deazaguanine, N⁴-ethanocytosine, N⁶-ethano-2,6-diaminopurine, 5-(C₃₋₆)-alkenyluracil, 5-(C₃₋₆)-alkynylcytosine, 5-fluorouracil and pseudocytosine.

5 Examples of useful intercalators are e.g. acridin, antraquinone, psoraten and pyrene.

Examples of useful nucleobase-binding groups are e.g. groups containing cyclic or heterocyclic rings. Non-limiting examples are 3-nitro pyrrole and 5-nitro indole.

- It is to be understood that alkyl, alkenyl and alkynyl groups may be branched or non-branched, cyclic or non-cyclic, and may further be interrupted by one or more heteroatoms, or may be unsubtituted or substituted by or may contain one or more functional groups. Non-limiting examples of such functional groups are acetyl groups, acyl groups, amino groups, carbamido groups, carbamoyl groups, carbamyl groups, carbonyl groups, carboxy groups, cyano groups, dithio groups, formyl groups, guanidino groups, halogens, hydrazino groups, hydrazo groups, hydroxamino groups, hydroxy groups, keto groups, mercapto groups, nitro groups, phospho groups, phosphon groups, phospho ester groups, sulfo groups, thiocyanato groups, cyclic, aromatic and heterocyclic groups.
- C₁₋₄ groups contain from 1 to 4 carbon atoms, C₁₋₆ groups contain from 1 to 6 carbon atoms, and C₁₋₁₅ groups contain from 1 to 15 carbon atoms, not including optional substituents, heteroatoms and/or functional groups. Non-limiting examples of such groups are -CH₃, -CF₃, -CH₂-, -CH₂CH₃, -CH₂CH₂-, -CH(CH₃)₂, -OCH₃, -OCH₂CH₃, -OCH₂CH₂-, -OCH(CH₃)₂, -OC(O)CH₃, -OC(O)CH₂-, -C(O)H, -C(O)-, -C(O)CH₃, -C(O)OH, -C(O)O-, -CH₂NH₂, -CH₂NH-, -CH₂OCH₃, -CH₂OCH₂-, -CH₂OC(O)OH, -CH₂OC(O)O-, -CH₂C(O)CH₂-, -C(O)NH₂, -CH=CH₂, -CH=CH-, -CH=CHCH₂C(O)OH, -CH=CHCH₂C(O)O-, -C=CH, -C=C-, -CH₂C=CH, -CH₂C=CH₃, -OCH₂C=CH, -OCH₂C=CCH₃, -NHCH₂C(O)-, -NHCH₂CH₂C(O)-, -NH(CH₂CH₂O)₂CH₂C(O)-, and HO(O)CCH₂C(O)(NH-(CH₂CH₂O)₂CH₂C(O))₂-, phenyl, benzyl, naphthyl, oxazolyl, pyridinyl, thiadiazolyl, triazolyl, and thienyl.

Within the present context, the expression "naturally occurring amino acid" is intended to comprise D- and L-forms of amino acids commonly found in nature, e.g. D- and L-forms of Ala (alanine), Arg (arginine), Asn (aspargine), Asp (aspartic acid), Cys (cysteine), Gln (glutamine), Glu (glutamic acid), His (histidine), Ile (isoleucine), Leu (leucine), Lys (lysine), Met (methionine), Phe (phenylalanine), Pro (proline), Ser (serine), Thr (threonine), Trp (tryptophan), Tyr (tyrosine) and Val (valine).

WO 98/15648 PCT/DK97/00425

In the present context, the expression "non-naturally occurring amino acid" is intended to comprise D- and L-forms of amino acids other than those commonly found in nature as well as modified naturally occurring amino acids. Examples of useful non-naturally occurring amino acids are D- and L-forms of β -Ala (β -alanine) Cha (cyclohexylalanine), Cit (citrulline), Hci (homocitrulline), HomoCys (homocystein), Hse (homoserine), Nle (norleucine), Nva (norvaline), Orn (ornithine), Sar (sarcosine) and Thi (thienylalanine).

In the present context, the term "sample" is intended to cover all types of samples suitable for the purpose of the invention. Examples of such samples are sputum, laryngeal swabs, gastric lavage, bronchial washings, biopsies, aspirates, expectorates, body fluids (spinal, pleural, pericardial, synovial, blood, pus, bone marrow), urine, tissue sections as well as food samples, soil, air and water samples. Analysis of samples originating from the before-mentioned samples (e.g. cultures and treated samples) are also within the scope of the invention.

10

20

25

30

35

In the present context, the term "hybrids" is intended to include complexes between a probe and a nucleic acid to be determined. Such hybrids may be made up of two or more strands.

The strength of the binding between the probe and the target nucleic acid sequence may be influenced by the ligand Q. When Q designates a nucleobase, Hoogsteen and/or Watson-Crick base pairing assist(s) in the formation of hybrids between a nucleic acid sequence to be detected and a probe. It is contemplated that one or more of the ligands may be a group which contribute little or none to the binding of the nucleic acid such as hydrogen. It is contemplated that suitable probes to be used comprise less than 25% by weight of peptide nucleic acid moieties, wherein Q designates such groups. One or more of the ligands Q may be groups that stabilise nucleobase stacking such as intercalators or nucleobase-binding groups.

In the above-indicated probes, one or more of the Q-groups may designate a label. Examples of suitable labels are given below. Moieties wherein Q denotes a label may preferably be located in one or both of the terminating moieties of the probe. Moieties wherein Q denotes a label may, however, also be located internally.

The peptide nucleic acid probes may comprise moieties, wherein all X groups are O (polyamides) or wherein all X groups are S (polythioamides). It is to be understood that the probes may also comprise mixed moieties (comprising both amide and thioamide moieties).

In another aspect, the present invention relates to peptide nucleic acid probes of formula (II), (III) and (IV) as well as mixtures of such probes defined in claim 15.

In a preferred embodiment, the peptide nucleic acid probes or mixtures thereof according to the invention are of formulas (I)-(IV) as defined in claim 16 with Z being NH, NCH₃ or O, each R², R³ and R⁴ independently being H or the side chain of a naturally occurring amino acid, the side chain of a non-naturally occurring amino acid, or C_{1.4} alkyl, and each Q being a naturally occurring nucleobase or a non-naturally occurring nucleobase with the provisos defined in claims 6 to 14.

Peptide nucleic acid probes or mixtures of such probes according to the invention are preferably those of formula (I)-(IV) as defined in claim 17 with Z being NH or O, and R² being H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, and Q being a nucleobase selected from thymine, adenine, cytosine, guanine, uracil, iso-C, and 2,6-diaminopurine with the provisos defined in claims 6 to 14.

Peptide nucleic acid probes or mixtures thereof, which are of major interest for detecting mycobacteria of the MTC group or one or more mycobacteria other than mycobacteria of the MTC group, are probes of formula (V) according to claim 18, wherein R⁴ is H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, Q is as defined in claim 17 and with the provisos indicated in claims 6 to 14.

The peptide nucleic acid probe comprises polymerised moieties as defined above and in the claims. From the formula, it is to be understood that the probe may comprise polymerised moieties which structure may be mutually different or identical. In some cases, it may be advantageous that at least one moiety of the probe, preferably one (or both) of the moieties terminating the probe, are of a different structure. Such terminating moieties may suitably be a moiety of formula (VI)

30

35

20

25

5

10

where Q is as defined above. Such moiety may suitably be connected to a peptide nucleic acid moiety through an amide bond.

The peptide nucleic acid probe according to the invention comprises from 6 to 30 polymerised moieties of formulas (I) to (V), and, in addition, optionally one or two terminating moieties of formula (VI) as defined above. The preferred length of the probe will depend on the sample material and whether labelled probes are used. It is contemplated that especially interesting probes comprise from 10 to 30 polymerised moieties of formulas (I) to (V), and, in addition,

WO 98/15648

5

10

15

20

optionally one or two terminating moieties of formula (VI) as defined above. Probes of the invention may suitably comprise from 12 to 25 polymerised moieties of formulas (I) to (V), more suitably from 14 to 22 polymerised moieties of formulas (I) to (V), most suitably from 15 to 20 polymerised moieties of formulas (I) to (V), and, in addition, optionally one or two terminating moieties of formula (VI).

As mentioned above, the polymerised moieties of the probes may be mutually different or identical. In some embodiments, the polymerised moieties of formulas (V) constitute at least 75% by weight (calculated by excluding labels and linkers), preferably at least 80% by weight and most preferably at least 90% by weight of the probe.

The ends on the moieties terminating the probe may be substituted by suitable substituents which in the following will be named "linkers". A terminating end may suitably be substituted by from 1 to 5 linkers, more suitably from 1 to 3 linkers. Such linkers may suitably be selected among C₁₋₁₅ alkyl, C₁₋₁₅ alkenyl and C₁₋₁₅ alkynyl groups as defined above. The linkers may be substituted or unsubstituted, branched or non-branched, or be interrupted by heteroatoms, or be substituted or contain functional groups as described above. This may depend on the chemical nature of the terminating moiety (i.e. whether the moiety is terminated by a carbon, oxygen or nitrogen atom). A terminating end or a linker on a terminating end may further be substituted by one or more labels, which labels may be incorporated end to end, i.e. so as to form a non-branched labelled end, or may be incorporated so as to form a branched labelled end ("zipper"). The linkers may be attached directly to a terminating end, may be attached to a label or between labels on a terminating end, or be attached to a terminating end before a label is attached to a terminating end. It should be understood that two terminating ends may carry different or identical substituents, linkers and/or labels. It should further be understood that the term "a label" is intended to comprise one or more labels as the term "linkers" is to comprise one or more linkers. For certain applications, it may be advantageous that one or more linkers are incorporated between the peptide nucleic acid moieties. Such applications may in particular be those based on triplex formation.

30

25

Examples of suitable linkers are -NH(CH₂CH₂O)_nCH₂C(O)-, -NH(CHOH)_nC(O)-, -(O)C(CH₂OCH₂)_nC(O)- and -NH(CH₂)_nC(O)-, NH₂(CH₂CH₂O)_nCH₂C(O)-, NH₂(CHOH)_nC(O)-, HO(O)C(CH₂OCH₂)_nC(O)-, NH₂(CH₂)_nC(O)-, -NH(CH₂CH₂O)_nCH₂C(O)OH, -NH(CHOH)_nC(O)OH, -(O)C(CH₂OCH₂)_nC(O)OH and -NH(CH₂)_nC(O)OH, wherein n is 0 or an integer from 1 to 8, preferably from 1 to 3. Examples of very interesting linkers are -NHCH₂C(O)-, -NHCH₂CH₂O(O)-, -NH(CH₂CH₂O)₂CH₂C(O)-, and HO(O)CCH₂CH₂C(O)(NH(CH₂CH₂O)₂CH₂C(O))₂-.

In the present context, the term "label" refers to a substituent which is useful for detection or visualisation. Suitable labels comprise fluorophores, biotin, dinitro benzoic acid, digoxigenin, radioisotope labels, peptide or enzyme labels, chemiluminiscence labels, fluorescent particles, hapten, antigen or antibody labels.

5

10

15

35

The expression "peptide label" is intended to mean a label comprising from 1 to 20 naturally occurring or non-naturally occurring amino acids, preferably from 1 to 10 naturally occurring or non-naturally occurring amino acids, more preferably from 1 to 8 naturally occurring or non-naturally occurring amino acids, most preferably from 1 to 4 naturally occurring or non-naturally occurring amino acids, linked together end to end in a non-branched or branched ("zipper") fashion. Such peptide label may be composed of amino acids which are mutually identical or different. In a preferred embodiment, such a non-branched or branched end comprises one or more, preferably from 1 to 8 labels, more preferably from 1 to 4, most preferably 1 or 2, further labels other than a peptide label. Such further labels may suitably terminate a non-branched end or a branched end. One or more linkers may suitably be attached to the terminating end before a peptide label and/or a further label. Furthermore, such peptide labels may be incorporated between the peptide nucleic acid moieties.

- The probe as such may also comprise one or more labels such as from 1 to 8, preferably from 1 to 4, most preferably 1 or 2, labels and/or one or more linker units, which may be attached internally, i.e. to the backbone of the probe. The linker units and labels may mutually be attached as described above.
- Examples of particular interesting labels are biotin, fluorescein labels, e.g. 5-(and 6)-carboxyfluorescein, 5- or 6-carboxyfluorescein, 6-(fluorescein)-5-(and 6)-carboxamido hexanoic acid
 and fluorescein isothiocyanate, peptide labels consisting of from 1 to 20 naturally occurring
 amino acids or non-naturally occurring amino acids, enzyme labels such as peroxidases like
 horse radish peroxidase (HRP), alkaline phosphatase, and soya bean peroxidase, dinitro
 benzoic acid, rhodamine, tetramethylrhodamine, cyanine dyes such as Cy2, Cy3 and Cy5,
 coumarin, R-phycoerythrin (RPE), allophycoerythrin, Texas Red, Princeton Red, and Oregon
 Green as well as conjugates of R-phycoerythrin and, e.g. Cy5 or Texas Red.

Examples of preferred labels are biotin, fluorescent labels, peptide labels, enzyme labels and dinitro benzoic acid. Peptide labels may preferably be composed of from 1 to 10, more preferably of from 1 to 8, most preferably of from 1 to 4, naturally occurring or non-naturally occurring amino acids. It may be particularly advantageous to incorporate one or more other labels as well as a peptide label such as from 1 to 8 or from 1 to 4 other labels, preferably 1 or

2 other labels.

20

25

30

Suitable peptide labels may preferably be composed of cysteine, glycine, lysine or ornithine.

- In a further embodiment, the Q substituent as defined above may be labelled. Suitable labels are as defined above. Between Q and such a label, a linker as defined above may be incorporated. It is preferred that such labelled ligands Q are selected from thymine and uracil labelled in the 5-position and 7-deazaguanine and 7-deazaguanine labelled in the 7-position.
- A mixture of peptide nucleic acid probes is also part of the present invention. Such mixture may comprise more than one probe capable of hybridising to 23S rRNA, and/or more than one probe capable of hybridising to 16S rRNA, and/or or more than one probe capable of hybridising to 5S rRNA. A mixture of probes may further comprise probe(s) directed to precursor rRNA and/or rDNA. The mixture may also comprise peptide nucleic acids for detecting more than one mycobacteria in the same assay.
 - In a preferred embodiment, the nucleobase sequence of the peptide nucleic acid probe is selected so as to be substantially complementary to the nucleobase sequence of the target sequence in question. In an especially preferred embodiment, the nucleobase sequence of the peptide nucleic acid probe is selected so as to be complementary to the nucleobase sequence of the target sequence in question. By "complementary" is meant that the nucleobases are selected so as to enable perfect match between the nucleobases of the probe and the nucleobases of the target, i.e. A to T or G to C. By substantially complementary is meant that the peptide nucleic acid probe is capable of hybridising to the target sequence, however, the probe does not necessarily have to be perfectly complementary to the target. For example, probes comprising one or more bases not complementary to the target sequence and nontarget sequences, especially base(s) located at the end of the probe, where the effect on the stability of probe-target nucleic acid hybrids is low. Another example is probes comprising other naturally occurring bases. Thus provided that the probe is capable of hybridising to the target sequence, the nucleobase difference(s) between target sequences and non-target sequences ensures that the stability of probe-non-target nucleic acid hybrids are lower than the stability of probe-target nucleic acid hybrids and therefore make such substantially complementary probes applicable for detection of mycobacteria.
- The probes may be synthesised according to the procedures described in "PNA Information Package" obtained from Millipore Corporation (Bedford, MA, USA), or may be synthesised on an Expedite Nucleic Acid Synthesis System (PerSeptive BioSystems, USA).

If using the Fmoc strategy for elongation of the probe with linkers or amino acids, it is possible to retain side chain amino groups protected with acid sensitive protection groups such as the Boc or Mtt group. This method allows introduction of a linker containing several Boc protected amino groups which can all be cleaved and labelled in the same synthesis cycle.

5

10

15

20

25

30

35

One way of labelling a probe is to use a fluorescent label, such as 5-(and 6)-carboxyfluorescein, 5- or 6-carboxyfluorescein, or 6-(fluorescein)-5-(and 6)-carboxamido hexanolc acid. The acid group is activated with HATU (O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) or HBTU (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) and reacted with the N-terminal amino group of the peptide nucleic acid. The same technique can be applied to other labelling groups containing an acid function. Alternatively, the succinimidyl ester of the above-mentioned labels may suitably be used or fluorescein isothiocyanate may be used directly.

After synthesis, probes can be cleaved from the resin using standard procedures as described by Millipore Corporation or PerSeptive BioSystems. The probes are subsequently purified and analysed using reversed-phase HPLC techniques at 50°C and were characterised by matrix-assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOFMS), plasma desorption mass spectrometry (PDMS), electron spray mass spectrometry (ESMS), or fast atom bombardment (FAB-MS).

Generally, probes such as probes comprising polymerised moieties of formula (IV) and (V) may also be prepared as described in, e.g., Angewandte Chemie, International Edition in English 35, 1939-1942 (1996) and Bioorganic & Medical Chemistry Letters, Vol 4, No 8, 1077-1080 (1994). Chemical properties of some probes are described in, e.g., Nature, 365, 566-568 (1993).

The method as claimed can be used for the detection of a target sequence of one or more mycobacteria optionally present in a sample. The method and the probes provide a valuable tool for analysing samples for the presence of such target sequences, hence providing information for establishing a diagnosis.

In the assay method according to the invention, the sample to be analysed for the presence of mycobacteria is brought into contact with one or more probes or a mixture of such probes according to the invention under conditions by which hybridisation between the probe(s) and any sample rRNA or rDNA originating from mycobacteria can occur, and the formed hybrids, if any, are observed or measured, and the observation or measurement is related to the presence of a target sequence of one or more mycobacteria. The observation or

measurement may be accomplished visually or by means of instrumentation.

centrifugation and decontamination methods mentioned above.

5

10

15

20

25

30

Prior to contact with probe(s) according to the invention, the samples may undergo various types of sample processing which include purification, decontamination and/or concentration. The sample may suitably be decontaminated by treatment with sodium hypochlorite and subsequently centrifuged for concentration of the mycobacteria. Samples e.g. sputum samples may be treated with a mucolytic agent such as N-Acetyl-L-cystein which reduces the viscosity of the sample as well as be treated with sodium hydroxide which decontaminates the sample, and subsequently centrifuged. Other well-known decontamination and concentration procedures include the Zephiran-trisodium phosphate method, Petroff's sodium hydroxide method, the oxalic acid method as well as the cetylpyridinium chloride-sodium chloride method. Samples may also be purified and concentrated by applying sample preparation methods such as filtration, immunocapture, two-phase separation either alone or in combination. The sample preparation methods may also be used together with the

Samples may, either directly or after having undergone one or more processing steps, be analysed in primarily two major types of assays, in situ-based assays and in vitro-based assays. In this context, in situ-based assays are to be understood as assays, in which the target nucleic acids remain within the bacterial cell during the hybridisation process. Examples are in situ hybridisation (ISH) assays on smears and biopsies as well as hybridisation to whole cells which may be in suspension and which subsequently may be analysed by e.g. flow cytometry optionally after capture of the bacteria onto particles (with same or different type and size), or by image analysis after spreading of the bacteria onto a solid medium, filter membrane or another substantially planar surface.

In vitro-based assays are to be understood as assays, in which the target nucleic acids are released from the bacterial cell before hybridisation. Examples of such assays are microtiter plate-based assays. Many other assay types, in which the released target nucleic acids by some means are captured onto a solid phase and subsequently analysed via a detector probe, are feasible and within the scope of the present invention. Even further, in vitro-based assays include assays, in which the target nucleic acids are not captured onto a solid phase, but in which the hybridisation and signal generation take place entirely in solution.

Samples for in situ-based assays may suitably be applied and optionally be immobilised to a support. Techniques for applying of a sample onto a solid support depend on the type of sample in question and include smearing and cytocentrifugation for liquid or liquified samples and sectioning of tissues for biopsy materials. The solid support may take a wide variety of

forms well-known in the art, such as a microscope slide, a filter membrane, a polymer membrane or a plate of various materials.

In the case of in vitro-based assays, the target nucleic acid may be released from the mycobacterial cells in various ways. Most methods for releasing the nucleic acids cause bursting of the cell wall (lysis) followed by extraction of the nucleic acids into a buffered solution. As mycobacteria have complex cell walls containing covalently associated peptidoglycans, arabinogalactans and in particular mycolic acids, they cannot easily be disrupted by standard methods used for the rapid lysis of other bacteria. Possible methods which are known to give successful lysis of the mycobacterial cell wall include methods which involve treatment with organic solvents, treatment with strong chaotropic reagents such as high concentrations of guanidine thiocyanate, enzyme treatment, bead beating, heat treatment, sonication and/or application of a French press.

Samples to be analysed by in situ assays may be fixed prior to hybridisation. The person skilled in the art will readily recognise that the appropriate procedure will depend on the type of sample to be examined. Fixation and/or immobilisation should preferably preserve the morphological integrity of the cellular matrix and of the nucleic acids. Examples of methods for fixation are flame fixation, heat fixation, chemical fixation and freezing. Flame fixation may be accomplished by passing the slide through the blue cone of a Bunsen burner 3 or 4 times; heat fixation may be accomplished by heating the sample to 80°C for 2 hours; chemical fixation may be accomplished by immersion of the sample in a fixative (e.g. formamide, methanol or ethanol). Freezing is particularly relevant for biopsies and tissue sections and is usually carried out in liquid nitrogen.

25

30

35

5

10

In one in situ hybridisation assay embodiment, the sample to be analysed is smeared onto a substantially planar solid support which may be a microscope slide, a filter membrane, a polymer membrane or another type of solid support with a planar surface. The preferred solid support is a microscope slide. After the smear has been prepared, it may optionally undergo further pre-treatment like treatment with bactericidal agents or additional fixation by immersion in e.g. ethanol. The sample may also be pre-treated with enzyme(s) which as primary function permeabilise the cells and/or reduce the viscosity of the sample. It may further be advantageous to perform a pre-hybridisation step in order to block sites which might otherwise give raise to non-specific binding. For this purpose, blocking agents like skim milk, and non-target probes may suitably be used. The components of the pre-hybridisation mixture should be selected so as to obtain an effective saturation of sites in the sample that might otherwise bind the probe non-specifically. The pre-hybridisation buffer may suitably comprise an appropriate buffer, blocking agent(s), and detergents.

During the in situ hybridisation, one or more probes according to the present invention are brought into contact with any target rRNA or rDNA inside the cells in a hybridisation solution under suitable stringency conditions. The concentration of the applied probe may vary depending on the chemical nature of the probe and the conditions under which hybridisation is carried out. Typically, a probe concentration between 1 nM and 1 µM is suitable. The hybridisation solution may comprise a denaturing agent which allows hybridisation to take place at a lower temperature than would be the case without the agent. The denaturing agent should be present in an amount effective to increase the ratio between specific binding and non-specific binding. The effective amount of denaturing agent depends on the type used and on the probe or combination of probes. Examples of denaturing agents are formamide, ethylene glycol and glycerol, and these may preferably be used in a concentration above 10% and less than 70%. The preferred denaturing agent is formamide which is used more preferably in concentrations from 20% to 60%, most preferably from 30% to 50%. It should be noted that in several instances it may not be necessary or advantageous to include a denaturing agent.

Prior to hybridisation or during hybridisation, a mixture of random probes (probes with random, non-selected sequences of optionally different length) may be added in excess to reduce non-specific binding. Also, one or more non-sense probes (probes with a defined nucleobase sequence and length differing from the nucleobase sequence of the target sequence) may be added in excess in order to reduce non-specific binding. Also, non-specific binding of detectable probes to one or more non-target nucleic acid sequences can be suppressed by addition of one or more unlabelled or independently detectable probes, which probes have a sequence that is complementary to the non-target sequence(s). It is particularly advantageous to add such blocking probes when the non-target sequence differs from the target sequence by only one mismatch.

It may be advantageous to include inert polymers which are believed to increase the effective concentration of the probe(s) in the hybridisation solution. One such macromolecule is dextran sulphate which may be used in concentrations of from 2.5% to 15%. The preferred concentration range is from 8% to 12% in the case of dextran sulphate. Other useful macromolecules are polyvinylpyrrolidone and ficoll, which typically are used at lower concentrations, e.g. 0.2%. It may further be advantageous to add one or more detergents which may reduce the degree of non-specific binding of the peptide nucleic acid probes. Examples of useful detergents are sodium dodecyl sulphate, Tween 20® or Triton X-100®. Detergents are usually used in concentrations between 0.05% and 1.0%, preferably between 0.05% and 0.25%. The hybridisation solution may furthermore contain salt. Although it is not

WO 98/15648 PCT/DK97/00425

necessary to include salt in order to obtain proper hybridisation, it may be advantageous to include salt in concentrations from 2 to 500 mM, or suitably from 5 to 100 mM.

During hybridisation, other important parameters are hybridisation temperature, concentration of the probe and hybridisation time. The person skilled in the art will readily recognise that optimal conditions must be determined for each of the above-mentioned parameters according to the specific situation, e.g. choice of probe(s) and type and concentration of the components of the hybridisation buffer, in particular the concentration of denaturing agent. Presence of volume excluders may also have an effect.

10

15

20

Following hybridisation, the sample is washed to remove any unbound and any non-specifically bound probe, and consequently, appropriate stringency conditions should be used. By stringency is meant the degree to which the reaction conditions favour the dissociation of the formed hybrids. The stringency may be increased typically by increasing the washing temperature and/or washing time. Typically, washing times from 5 to 40 minutes may be sufficient. Two or more washing periods of shorter time may also give good results. A range of buffers may be used, including biological buffers, phosphate buffers and standard citrate buffers. The demand for low salt concentration in the buffers is not as pertinent as for DNA probe assays due to the difference response to salt concentration. In some cases, it is advantageous to increase the pH of the washing buffer as it may give an increased signal-to noise ratio (see WO 97/18325). This is conceivably due to a significant reduction of the non-specific binding. Thus, it may be advantageous to use a washing solution with a pH value form 8 to 10.5, preferably from 9 to 10.

Visualisation of bound probe(s) must be carried out with due regard to the type of label chosen. There are a wide range of useful probe labels, in particular various fluorescent labels such as fluorescein, rhodamine and derivatives thereof. Furthermore, labels like enzymes (e.g. peroxidases and phosphatases) and haptens (e.g. biotin, digoxigenin, dinitro benzoic acid) may suitably be applied. In the case of fluorescent labels, the hybrids formed may be visualised using a microscope with a magnification of at least × 250, preferably × 1000. The visualisation may further be carried out using a CCD (charge coupled device) camera optionally controlled by a computer. When haptens are used as labels, the hybrids may be detected using an antibody conjugated with an enzyme. In these cases, the detection step

may be based on colorimetry, fluorescence or luminescence.

35

The probes may alternatively be labelled with fluorescent particles having the fluorescent label embedded in the particles (e.g. Estapor K coloured microspheres), located on the surface of the particles and/or coupled to the surfaces of the particles. As the particles have to come into

contact with the target nucleic acids within the cells, the use of fluorescent particles may necessitate pretreatment of the bacteria. Relatively small particles e.g. about 20 nm may suitable be used.

In another in situ hybridisation embodiment, frozen tissue or biopsy samples are cut into thin 5 sections and transferred to a substantially planar surface, preferably microscope slides. Prior to hybridisation, the tissue or biopsy may be treated with a fixative, preferably a precipitating fixative such as acetone, or the sample is incubated in a solution of buffered formaldehyde. Alternatively, the biopsy or tissue section can be transferred to a fixative such as buffered formaldehyde for 12 to 24 hours and following fixation, the tissue may be embedded in paraffin 10 forming a block from which thin sections can be cut. Prior to hybridisation, the tissue section is dewaxed and rehydrated using standard procedures. Permeabilisation (e.g. treatment with proteases, diluted acids, detergents, alcohol and/or heat) may in some cases be advantageous. The selected method for permeabilisation depends on several factors, for 15 instance on the fixative used, the extent of fixation, the type and size of sample, and on the applied probe. For these types of samples, sample processing, prehybridisation, hybridisation, washing and visualisation may be carried out using same or adjusted conditions as described above.

In a further embodiment of the in situ assays, the bacterial cells are kept in suspension during fixation, prehybridisation, hybridisation and washing are carried out under the same or similar conditions as described above. The preferred type of label for this embodiment is fluorescent labels. This allows detection of hybridised cells by flow cytometry, recording the intensity of fluorescence per cell. Bacterial cells in suspension may further be coupled to particles, preferably with a size of from 20 nm to 10 µm. The particles may be made of materials well-known in the art like latex, dextran, cellulose and/or agarose, and may optionally be paramagnetic or contain a fluorescent label. Normally, bacterial cells are coupled to particles using antibodies against the target bacteria, but other means like molecular imprinting may also be used. Coupling of the bacterial cells to particles may be advantageous in sample handling and/or during detection.

20

25

30

35

In the embodiments of in situ hybridisation described above, the probes according to the invention are used for detecting a target sequence of one or more mycobacteria. In a preferred embodiment, the probes are suitable for detecting a target sequence of mycobacteria of the Mycobacterium tuberculosis Complex (MTC), mycobacteria other than the Mycobacterium tuberculosis Complex (MOTT), or mycobacteria of the Mycobacterium avium Complex (MAC). The probes are further suitable for detecting simultaneously different target sequences originating from the same mycobacteria.

5

10

15

20

Samples to be analysed using in vitro-based assays need to undergo a treatment by which the nucleic acids are released from the bacterial cells. Nucleic acids may be released using organic solvents, strong chaotropic reagents such as high concentrations of guanidine thiocyanate, enzymes, bead beating, heating, sonication and/or application of a French press. The obtained nucleic acids may undergo additional purification prior to hybridisation.

In one in vitro hybridisation embodiment, the sample comprising the target nucleic acid is added to a container comprising immobilised capture probe(s) and one or more probe(s) labelled to function as detector probe(s). The hybridisation should be performed under suitable stringency conditions. The hybridisation solution may further comprise a denaturing agent, blocking probes, inert polymers, detergents and salt as described for the in situ-type assays. Likewise, the hybridisation temperature, probe concentration and hybridisation time are important parameters that need to be controlled according to the specific conditions of the assay, e.g. choice of peptide nucleic acid probe(s) and concentration of some of the ingredients of the hybridisation buffer. If hybridisation of the target nucleic acid to the capture probe(s) and detector probe(s), respectively, is performed in two separate steps, different parameters, in particular different stringency conditions, may be used in these steps. The concentration of the capture probe may be higher for in situ assays as hybridisation may be controlled better and washing can be performed more efficiently.

The capture probes may be immobilised onto a solid support by any means, e.g. by a coupling reaction between a carboxylic acid on a linker and an amino derivatised support. The capture probe may further be coupled onto the solid support by photochemical activation of photoreactive groups which have been attached absorptively to the solid support prior to photochemical activation. Such photoreactive groups are described in the US 5 316 784 A. The capture probes may further be coupled to a hapten which allows an affinity based immobilisation to the solid support. One such example is coupling of a biotin to the probe(s) and immobilisation via binding to a steptavidin-coated surface.

30

35

25

The solid support may take a wide variety of forms well-known in the art, such as a microtiter plate having one or more wells, a filter membrane, a polymer membrane, a tube, a dip stick, a strip and particles. Filter membranes may be made of cellulose, celluloseacetate, polyvinylidene fluoride or any other materials well-known in the art. The polymer membranes may be of polystyrene, nylon, polypropylene or any other materials well known in the art. Particles may be paramagnetic beads, beads made of polystyrene, polypropylene, polyethylene, dextran, nylon, amyloses, celluloses, polyacrylamides and agarose. When the solid support has the form of a filter, a membrane, a strip or beads, it (they) may be

incorporated into a single-use device.

The selection of the label of the detector probe(s) depend on the specific assay format and possible instrumentation. When biotin labelled probes are used, the hybrids may be detected using streptavidin or an antibody against the biotin label which antibody or streptavidin may be conjugated with an enzyme and the actual detection depend on the choice of the specific enzyme, preferably a phosphatase or a peroxidase, and the substrate for the selected enzyme. The signal may in some cases be enhanced using commercially available amplification systems such as the catalysed signal amplification system for biotinylates probes (CSA by DAKO). Various polymer-based enhancement systems may also be used. An example is a dextran polymer to which both a hapten specific antibody and an enzyme is coupled. The detector probe(s) may further be labelled with other haptens, e.g. digoxigenin, dinitro benzoic acid and fluorescein, in which case the hybrids may be detected using an antibody against the hapten which antibody may be conjugated with an enzyme. It is even possible to apply detector probe(s) which have enzymes coupled directly onto the probes. There are a wide range of possibilities for selection of enzyme substrates allowing for colourimetric (substrates e.g. p-nitro-phenyl phosphate or tetra-methyl-benzidine), fluorogenic (substrates e.g. 4-methylumbilliferylphosphate) or chemiluminescent (substrates e.g. 1,2dioxetanes) detection.

20

25

30

10

15

The detector probes may further be labelled with various fluorescent labels, preferably fluorescein or rhodamine, in which case the hybrids may be detected by measuring the fluorescence.

The detector probe(s) will typically be different from the capture probe(s), thus ensuring dual species specificity. The dual specificity will most often allow at least one of the probes to be shorter, e.g. a 10 mer probe.

Furthermore, the capture of purine rich sequences may be improved by utilising bis-peptide nucleic acids as capture probes. Such bis-peptide nucleic acids are described in WO 96/02558. The bis-peptide nucleic acids comprise a first peptide nucleic acid strand capable of hybridising in parallel fashion to the target nucleic acid, and a second peptide nucleic acid strand capable of hybridising in antiparallel fashion to the purine rich sequence of the nucleic acid to be captured. The two peptide nucleic acid strands are connected by a linker and are in this way capable of forming a triplex structure with said purine rich sequence nucleic acid. The number of polymerised moieties of each linker-separated peptide nucleic acid may be as previously defined for non-bis-peptide nucleic acids. However, due to the high stability of the triplexes formed, bis-peptide nucleic acids with short first and second strands can be used

making the design of a pyrimidine rich probe easier.

Instead of using a detector probe, capture probe: nucleic acid complexes may be detected using a detection system based on an antibody reacting specifically with complexes formed between peptide nucleic acids and nucleic acids (such as described in WO 95/17430), in which detection system the primary antibody may comprise a label, or which detection system comprises a labelled secondary antibody, which specifically binds to the primary antibody. The specific detection again depends on the selected substrate which may be of any type of those mentioned above.

10

5

Depending on the type of specific assay format, label and detection principle various types of instrumentation may be used including conventional microplate readers, luminometers and flow cytometers. Adaptation of adequate instrumentation may allow for automatisation of the assay.

15

25

30

In an example of this embodiment, a capture probe of the present invention is coupled to a microtiter plate by a photochemical reaction between antraquinon-labelled capture probe and polystyrene of the microwell. Target rRNA is added to the microwells and incubated under stringent conditions. Unbound rRNA is removed by washing and the microwell are incubated with a hapten-labelled detector probe under stringent conditions. The visualisation is carried out using an enzyme-labelled antibody against the hapten, which after removal of unbound antibody is detected using a chemiluminescence substrate.

In another example of this embodiment capture probes are coupled to latex particles, and hybridisation is carried out under suitable conditions in the presence of e.g. fluorescein labelled detector probe(s). After hybridisation and optionally washing, the hybrids are detected by flow cytometry. A range of different beads (e.g. by size or colours) may carry different capture probes for different targets, thus allowing a multiple detection system.

In a further embodiment of the in vitro assays format, the capture probe, the target nucleic acid and the detector probe may hybridise in solution, and subsequently the capture probe is attached to a solid phase. The solid phase, the hybridisation conditions and means of detection may be selected according to the specific method as described above.

In a further embodiment of in vitro assays, the target nucleic acid may be immobilised onto filter or polymer membranes or other types of solid phases well-known in the art. The hybridisation conditions and means of detection may be selected according to the specific setup as described above.

In a further embodiment of the in vitro assay, an array of up to 100 or even more different probes directed against different target sequences may be immobilised onto a solid surface and hybridisation of the target sequences to all the probes is carried out simultaneously. The solid phase, the hybridisation conditions and means of detection may be as described above. This allow for simultaneous detection or identification of a range of parameters, i.e. species identification and resistance patterns.

The present probes further provide a method of diagnosing infection by mycobacteria and a method for determining the stage of the infection and the appropriate treatment by which methods one or more optionally labelled probes according to the invention are brought into contact with a patient sample and the type of treatment and/or the effect of a treatment is (are) evaluated.

Kits comprising at least one peptide nucleic acid probe as defined herein are also part of the present invention. Such kit may further comprise a detection system with at least one detecting reagent and/or a solid phase capture system.

DESCRIPTION OF SPECIFIC EMBODIMENTS

20

25

30

10

Examples of suitable Qs of adjacent moieties are given below. Peptide nucleic acid probes comprising such Qs will be suitable for detecting mycobacteria, in particular mycobacteria of the MTC group or mycobacteria other than mycobacteria of the MTC group. The probes are written from left to right corresponding to from the N-terminal end towards the C-terminal end. Suitable Q subsequences for detecting 23S and 16S rRNA as well as 5S rRNA of the MTC group are given below. Suitable Q subsequences for detecting 23S and 16S rRNA of mycobacteria other than mycobacteria of the MTC group are further given below. The Q subsequences include at least one nucleobase complementary to a nucleobase selected from the positions given in parenthesis. The Q subsequences are given as non-limiting examples of construction of suitable probe nucleobase sequences. It is to be understood that the probes may comprise fewer or more peptide nucleic acid moieties than indicated.

MTC group (23S)

	AGA TGC GGG TAG CAC (selected from positions 149-158 in Figure 1A),	(Seq ID no 1)
35	TGT TTT CTC CTC CTA (selected from positions 220-221 in Figure 1A),	(Seq ID no 2)
	ACT GCC TCT CAG CCG (selected from positions 328-361 in	
	Figure 1A and Figure 1B),	(Seq ID no 3)
	TGA TAC TAG GCA GGT (selected from positions 453-455 in Figure 1B),	(Seq ID no 4)
	CGG ATT CAC AGC GGA (selected from positions 490-501 in Figure 1B),	(Seg ID no 5)

	TCA CCA CCC TCC TCC (selected from positions 637-660 in Figure 1C),	(Seq ID no 6)
	CCA CCC TCC (selected from positions 637-660 in Figure 1C)	(modified Seq ID no 6)
	TTA ACC TTG CGA CAT (selected from positions 706-712 in Figure 1D),	(Seq ID no 7)
	ACT ATT CAC ACG CGC (selected from positions 762-789 in Figure 1D),	(Seq ID no 8)
. 5	CTC CGC GGT GAA CCA (selected from position 989 in Figure 1D),	(Seq ID no 9)
	GCT TTA CAC CAC GGC (selected from positions 1068-1072 in Figure 1E),	(Seq ID no 10)
	ACG CTT GGG GGC CTT (selected from position 1148 in Figure 1E),	(Seq ID no 11)
	CCA CAC CCA CCA CAA (selected from positions 1311-1329 In Figure 1E),	(Seq ID no 12)
	CCG GTG GCT TCG CTG (selected from positions 1361-1364 in Figure 1F),	(Seq ID no 13)
10	ACT TGC CTT GTC GCT (selected from position 1418 in Figure 1F),	(Seq ID no 14)
	GAT TCG TCA CGG GCG (selected from positions 1563-1570 in Figure 1F),	(Seq ID no 15)
	AAC TCC ACA CCC CCG (selected from positions 1627-1638 in Figure 1G),	(Seq ID no 16)
	ACT CCA CAC CCC CGA (selected from positions 1627-1638 in Figure 1G),	(Seq ID no 17)
	ACC CCT TCG CTT GAC (selected from positions 1675-1677 in Figure 1G),	(Seq ID no 18)
15	CTT GCC CCA GTG TTA (selected from position 1718 in Figure 1G),	(Seq ID no 19)
	CTC TCC CTA CCG GCT (selected from positions 1734-1740 in Figure 1H),	(Seq ID no 20)
	GAT ATT CCG GTC CCC (selected from positions 1967-1976 in Figure 1H),	(Seq ID no 21)
	ACT CCG CCC CAA CTG (selected from positions 2403-2420 in Figure 1H).	(Seq ID no 22)
	CTG TCC CTA AAC CCG (selected from positions 2457-2488 in Figure 1I),	(Seq ID no 23)
- 20	TTC GAG GTT AGA TGC (selected from positions 2457-2488 in Figure 1I).	(Seq ID no 24)
	GTC CCT AAA CCC GAT (selected from positions 2457-2488 in Figure 1I),	(Seq ID no 25)
	GGT GCA CCA GAG GTT (selected from positions 2952-2956 in Figure 1I),	(Seq ID no 26)
	CTG GCG GGA CAA CTG (selected from positions 2966-2969 in Figure 1J),	(Seq ID no 27)
	TTA TCC TGA CCG AAC (selected from positions 3000-3003 in Figure 1J),	(Seq ID no 28)
25	GAC CTA TTG AAC CCG (selected from positions 3097-3106 in Figure 1J),	(Seq ID no 29)
	MTC group (16S)	
	GAA GAG ACC TTT CCG (selected from positions 76-79 in Figure 2A).	(Seq ID no 30)
	CAC TCG AGT ATC TCC (selected from positions 98-101 in Figure 2A),	(Seg ID no 31)
30	ATC ACC CAC GTG TTA (selected from positions 138-136 in Figure 2A),	(Seq ID no 32)
	GCA TCC CGT GGT CCT (selected from positions 194-201 in Figure 2B),	(Seq ID no 33)
	CAC AAG ACA TGC ATC (selected from positions 194-201 in Figure 2B),	(Seq ID no 34)
	TAA AGC GCT TTC CAC (selected from positions 222-229 in Figure 2B),	(Seq ID no 35)
	GCT CAT CCC ACA CCG (selected from position 242 in Figure 2B),	(Seq ID no 36)
35	CCG AGA GAA CCC GGA (selected from position 474 in Figure 2C),	(Seq ID no 37)
	AGT CCC CAC CAT TAC (selected from positions 1136-1145 in Figure 2C),	(Seq ID no 38)
	AAC CTC GCG GCA TCG (selected from positions 1271-1272 in Figure 2C),	(Seq ID no 39)
	GGC TTT TAA GGA TTC (selected from positions 1287-1292 in Figure 2D),	(Seq ID no 40)
	GAC CCC GAT CCG AAC (selected from position 1313 in Figure 2D),	(Seq ID no 41)
40	CCG ACT TCA CGG GGT (selected from position 1334 in Figure 2D),	(Seq ID no 42)

MTC group (5S)

	CGG AGG GGC AGT ATC (selected from positions 86-90 in Figure 3),	(Seq ID no 43)
	Mycobacteria other than those of the MTC group (23S)	
	GAT CAA TGC TCG GTT (selected from positions 99-101 in Figure 4A),	(Seq ID no 44)
5	TTC CCC GCG TTA CCT (selected from position 183 in Figure 4A),	(Seq ID no 45)
	TTA GCC TGT TCC GGT (selected from positions 261-271 in Figure 4A),	(Seq ID no 46)
	GCA TGC GGT TTA GCC (selected from positions 281-284 in Figure 4B),	(Seq ID no 47)
	TAC CCG GTT GTC CAT (selected from positions 290-293 in Figure 4B),	(Seq ID no 48)
	GTA GAG CTG AGA CAT (selected from positions 327-335 and	,
10	343-357 in Figure 4B),	(Seq ID no 49)
	GCC GTC CCA GGC CAC (selected from positions 400-405 in	, ,
	Figure 4B and Figure 4C),	(Seq ID no 50)
	CTC GGG TGT TGA TAT (selected from positions 453-462 in Figure 4C),	(Seq ID no 51)
	ACT ATT TCA CTC CCT (selected from positions 587-599 in Figure 4C),	(Seq ID no 52)
15	ACG CCA TCA CCC CAC (selected from positions 637-660 in Figure 4D),	(Seq ID no 53)
	CGA CGT GTC CCT GAC (selected from positions 704-712 in Figure 4D),	(Seq ID no 54)
	ACT ACA CCC CAA AGG (selected from positions 763-789 in Figure 4E),	(Seq ID no 55)
	CAC GCT TTT ACA CCA (selected from positions 1060-1074 in Figure 4E),	(Seq ID no 56)
	GCG ACT ACA CAT CCT (selected from positions 1177-1185 in Figure 4E),	(Seq ID no 57)
20	CGG CGC ATA ATC ACT (selected from positions 1259-1265 in Figure 4E),	(Seq ID no 58)
	CCA CAT CCA CCG TAA (selected from positions 1311-1327 in Figure 4F),	(Seq ID no 59)
	CGC TGA ATG GGG GAC (selected from positions 1345-1348 in Figure 4F),	(Seq ID no 60)
	GGA GCT TCG CTG AAT (selected from positions 1361-1364 in Figure 4G),	(Seq ID no 61)
-	CGG TCA CCC GGA GCT (selected from positions 1361-1364 in Figure 4G),	(Seq ID no 62)
25	GGA CGC CCA TAC ACG (selected from positions 1556-1570 in Figure 4G),	(Seq ID no 63)
	GAA GGG GAA TGG TCG (selected from positions 1608-1613 in Figure 4H),	(Seq ID no 64)
	AAT CGC CAC GCC CCC (selected from positions 1626-1638 in Figure 4H),	(Seq ID no 65)
	CAG CGA AGG TCC CAC (selected from positions 1651-1659 in Figure 4H),	(Seq ID no 66)
	GTC ACC CCA TTG CTT (selected from positions 1675-1677 in Figure 4H),	(Seq ID no 67)
30	ATC GCT CTC TAC GGG (selected from positions 1734-1741 in Figure 4H),	(Seq ID no 68)
	GTG TAT GTG CTC GCT (selected from positions 1847-1853 in Figure 4I),	(Seq ID no 69)
	ACG GTA TTC CGG GCC (selected from positions 1967-1976 in Figure 4!),	(Seq ID no 70)
	GGC CGA ATC CCG CTC (selected from positions 2006-2010 in Figure 4I),	(Seq ID no 71)
	AAA CAG TCG CTA CCC (selected from positions 2025-2027 in Figure 4I),	(Seq ID no 72)
35	CCT TAC GGG TTA ACG (selected from positions 2131-2132 in Figure 4J),	(Seq ID no 73)
	GAG ACA GTT GGG AAG (selected from positions 2252-2255 in Figure 4J),	(Seq ID no 74)
	TGG CGT CTG TGC TTC (selected from positions 2396-2405 in	
	Figure 4J and Figure 4K),	(Seq ID no 75)
40	CGA CTC CAC ACA AAC (selected from positions 2416-2420 in Figure 4K),	(Seq ID no 76)
40	GAT AAG GGT TCG ACG (selected from positions 2474-2478 in Figure 4K),	(Seq ID no 77)
	ATC CGT TGA GTG ACA (selected from position 2687 in Figure 4K),	(Seq ID no 78)
	CAG CCC GTT ATC CCC (selected from position 2719 in Figure 4K),	(Seq ID no 79)

DO	TAL	(97/N	0.43E
-	1/11/20	. 4 //10	14 Z.3

WO 98/15648

35

TTC CTT TTA GTT TTA (selected from positions 865 in Figure 7),	(Seq ID no 113)
TTC CTT AGA GTT TTA (selected from positions 866 in Figure 7).	(Seq ID no 114)
TTC CTT CGA GTT TTA (selected from positions 866 in Figure 7),	(Seq ID no 115)
TTC CTT GGA GTT TTA (selected from positions 866 in Figure 7),	(Seq ID no 116)

5

Other examples of suitable Q subsequences are given below.

CAT GTG TCC TGT GGT and (Seq ID no 117)
CGT CAG CCC GAG AAA (Seq ID no 118)

selected so as to be complementary to M. gordonae 16S rRNA (positions 174-188 and 452-466, respectively, of GenBank entry GB:MSGRR16SI, accession no. M29563). These positions correspond to positions 192-206 and 473-487, respectively, of the alignments shown in Figure 2 and 5. Probes having this or a similar nucleobase sequence are suitable for detecting M. gordonae.

15

20

CAC TAC ACA CGC TCG, and

(Seq ID no 119)
TGG CGT TGA GGT TTC

(Seq ID no 120)
selected so as to be complementary to positions 781-795 and 2369-2383, respectively, of M.
kansasii 23S rRNA (GenBank entry MK23SRRNA accession number Z17212). These
positions correspond to positions 774-794 and 2398-2412, respectively, of the alignments
shown in Figure 1 and 4. Probes having this or a similar nucleobase sequence are suitable for

Precursor rRNA

25 AAC ACT CCC TTT GGA

detecting M. kansasii.

(Seq ID no 123)

A peptide nucleic acid probe having the above-indicated nucleobase sequence is directed to M. tuberculosis precursor rRNA. The probe is complementary to positions 602 to 616 of GenBank accession number X58890.

30

35

Especially, probes based on those nucleobase sequences with sequence identification numbers Seq ID no 62, 79 and 80 (and other probes selected from positions 1361-1364 in Figure 1F, 2719 in Figure 4K and 2809 in Figure 4L) are suitable for detecting M. avium. Probes based on the nucleobase sequence with sequence identification number Seq ID no 55 (and other probes selected from positions 763-789 in Figure 4E) are suitable for detecting M. avium, M. intracellulare and M. scrofulaceum as a group (the organisms termed the MAIS group of mycobacteria). In addition, probes based on the nucleobase sequences with sequence identification numbers Seq ID no 77 and 81 are suitable for detecting M. avium, M. intracellulare and M. paratuberculosis as a group.

The invention is further illustrated by the non-limiting examples given below.

EXAMPLES

5

10

15

EXAMPLE 1

Mycobacterium species (M. bovis and M. intracellulare) 23S rDNA were partly amplified by PCR, and the PCR products were sequenced (both strands) using Cy5-labelled oligonucleotide primers (DNA Technology, Aarhus, Denmark) and the 7-deaza-dGTP Thermo Sequenase cycle sequencing kit from Amersham, Little Chalfont, England. Sequences were read using an ALFexpress automated sequencer and ALFwin (version 1.10) software from Pharmacia Biotech, Uppsala, Sweden. M. bovis and M. intracellulare 23S rRNA sequences are included at the following positions of the 23S rDNA sequence alignments: positions 681-729 (Figures 1C and 4D), positions 761-800 (Figures 1D and 4E), positions 2401-2440 (Figures 1H and 4K), positions 2441-2480 (Figures 1I and 4K), positions 2481-2520 (Figure 1I), positions 3041-3080 (Figure 4L), and positions 3081-3120 (Figures 1J and 4L).

EXAMPLE 2

20

Sequence alignments (see Figures 1 to 5) of 23S, 16S and 5S rDNA of mycobacteria of the MTC group, and 23S and 16S rDNA of mycobacteria other than those of the MTC group (MOTT) were done using the Megalign (version 3.12) alignment tool from DNASTAR (Madison, WI, USA). Up to one hundred sequences were aligned at a time.

25

30

Peptide nucleic acid probes in which the nucleobase sequence was complementary to distinctive mycobacterial rRNA were designed with due regard to secondary structures using the PrimerSelect program (version 3.04) from DNASTAR. As a control of sequence specificity, all probe sequences were subsequently matched with the GenBank and EMBL databases using BLAST sequence similarity searching at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov).

As examples, the following sequences were selected:

35 MTC 23S

TCA CCA CCC TCC TCC
CCA CCC TCC TCC
ACT ATT CAC ACG CGC
CCA CAC CCA CCA CAA

(Seq ID no 6) (modified Seq ID no 6) (Seq ID no 8) (Seq ID no 12)

	WO 98/15648	PCT/DK97/004	25
		37	
	AAC TCC ACA CCC CCG	(Seq ID no 16)	
	ACT CCA CAC CCC CGA	(Seq ID no 17)	
	ACT CCG CCC CAA CTG	(Seq ID no 22)	
	CTG TCC CTA AAC CCG	(Seq ID no 23)	
5	TTC GAG GTT AGA TGC	(Seq ID no 24)	
	GTC CCT AAA CCC GAT	(Seq ID no 25)	
	GAC CTA TTG AAC CCG	(Seq ID no 29)	
	MTC 16S	·	
10	GCA TCC CGT GGT CCT	(Seq ID no 33)	
	CAC AAG ACA TGC ATC	(Seq ID no 34)	
	GGC TTT TAA GGA TTC	(Seq ID no 40)	
	MOTT 23S	·	
15	GAT CAA TGC TCG GTT	(Seq ID no 44)	
	CGA CTC CAC ACA AAC	(Seq ID no 76)	
	MOTT 16S		
	GCA TTA CCC GCT GGC	(Con ID no OE)	
20		(Seq ID no 85)	
	Drug resistance		
	GTC TTA TCG TCC TGC	(Seq ID no 90)	
	GTC TTC TCG TCC TGC	(Seq ID no 91)	
	GTC TTG TCG TCC TGC	(Seq iD no 92)	
25	GTC TAT TCG TCC TGC	(Seq ID no 93)	
•	GTC TCT TCG TCC TGC	(Seq ID no 94)	
	GTC TGT TCG TCC TGC	(Seq ID no 95)	
	Precursor rRNA		
30	AAC ACT CCC TTT GGA	(Seq ID no 123)	
	Non-sense probes		
	GTC CGT GAA CCC GAT	/See ID == 404)	
	TAC GCT CTT TGA GCT	(Seq ID no 121) (Seq ID no 122)	
35	,	(Sey ID NO 122)	
	EXAMPLE 3	•	

Peptide nucleic acid probes were synthesised using an Expedite 8909 Nucleic Acid Synthesis System purchased from PerSeptive Biosystems (Framingham, USA). The peptide nucleic acid probes were terminated with two β -alanine molecules or with one or two lysine molecule(s) and, before cleavage from the resin, labelled with 5-(or 6)-carboxyfluorescein (Flu) or

(OK 745/modified Seq ID no 89)

(OK 746/modified Seq ID no 90)

rhodamine (Rho) at the β -amino group of alanine (peptide label) or ϵ -amino group of lysine (peptide label), respectively. Probes were purified using reverse phase HPLC at 50°C and characterised using a G2025 A MALDI-TOF MS instrument (Hewlett Packard, San Fernando, California, USA). Molecular weights determined were within 0.1% of the calculated molecular weights.

The following labelled peptide nucleic acid probes were synthesised:

MTC 23S

40

Lys(Flu)-GTC TTT TCG TCC TGC-NH2

Lys(Rho)-GTC TTA TCG TCC TGC-NH2

	MTC 23S	
10	Lys(Flu)-Lys(Flu)-TCA CCA CCC TCC TCC-NH2	(OK 446/modified Seq ID no 6)
	Lys(Flu)-Lys(Flu)-CCA CCC TCC TCC-NH ₂	(OK 575/modified Seq ID no 6)
	Lys(Flu)-Lys(Flu)-ACT ATT CAC ACG CGC-NH ₂	(OK 447/modified Seq ID no 8)
	Lys(Flu)-ACT ATT CAC ACG CGC-NH ₂	(OK 688/modified Seq ID no 8)
	Lyś(Fiu)-Lyś(Fiu)-CCA CAC CCA CCA CAA-NH ₂	(OK 448/modified Seq ID no 12)
15	Lys(Flu)-Lys(Flu)-AAC TCC ACA CCC CCG-NH2	(OK 449/modified Seq ID no 16)
	Lys(Flu)-Lys(Flu)-ACT CCA CAC CCC CGA-NH ₂	(OK 309/modified Seq ID no 17)
	Lys(Flu)-Lys(Flu)-ACT CCG CCC CAA CTG-NH₂	(OK 450/modified Seq ID no 22)
	Lys(Flu)-Lys(Flu)-CTG TCC CTA AAC CCG-NH2	(OK 305/modified Seq ID no 23)
	Lys(Flu)-Lys(Flu)-TTC GAG GTT AGA TGC-NH₂	(OK 306/modified Seq ID no 24)
20	Lys(Flu)-TTC GAG GTT AGA TGC-NH2	(OK 682/modified Seq ID no 24)
	Lys(Flu)-Lys(Flu)-GTC CCT AAA CCC GAT-NH2	(OK 307/modified Seq ID no 25)
	Lys(Flu)-GTC CCT AAA CCC GAT-NH2	(OK 654/modified Seq ID no 25)
	Lys(Flu)-GAC CTA TTG AAC CCG-NH ₂	(OK 660/modified Seq ID no 29)
25	MTC 16S	
	Lys(Flu)-Lys(Flu)-Gly-GCA TCC CGT GGT CCT-NH₂	(OK 223/modified Seq ID no 33)
	Lys(Flu)-Lys(Flu)-CAC AAG ACA TGC ATC-NH2	(OK 310/modified Seq ID no 34)
	Lys(Flu)-CAC AAG ACA TGC ATC-NH2	(OK 655/modified Seq ID no 34)
	Lys(Flu)-GGC TTT TAA GGA TTC-NH2	(OK 689/modified Seq ID no 40)
30	Lys(Rho)-GGC TTT TAA GGA TTC-NH2	(OK 702/modified Seq ID no 40)
	MOTT 23S	
	Flu-β-Ala-β-Ala-GAT CAA TGC TCG GTT-NH ₂	(OK 624/modified Seq ID no 44)
	Flu-β-Ala-β-Ala-CGA CTC CAC ACA AAC-NH ₂	(OK 612/modified Seq ID no 76)
35	MOTT 16S	
	Flu-β-Ala-β-Ala-GCA TTA CCC GCT GGC-NH ₂	(OK 623/modified Seq ID no 85)
	,	(On ozomodnied Sed in 10 82)
	Drug resistance	

WO 98/15648

39

PCT/DK97/00425

Lys(Rho)-GTC TTC TCG TCC TGC-NH ₂
Lys(Rho)-GTC TTG TCG TCC TGC-NH ₂
Lys(Rho)-GTC TAT TCG TCC TGC-NH ₂
Lys(Rho)-GTC TCT TCG TCC TGC-NH ₂
Lys(Rho)-GTC TGT TCG TCC TGC-NH ₂

(OK 746/modified Seq ID no 91) (OK 746/modified Seq ID no 92) (OK 747/modified Seq ID no 93) (OK 747/modified Seq ID no 94) (OK 747/modified Seq ID no 95)

Precursor rRNA

Lys(Fiu)-AAC ACT CCC TTT GGA-NH,

(OK 749/modified Seq ID no 123)

10 Reduction of non-specific binding

GTC CGT GAA CCC GAT-NH₂
Gly-TAC GCT CTT TGA GCT-NH₂

(OK 507/modified Seq ID no 121) (OK 714/modified Seq ID no 122)

EXAMPLE 4

15

Initially the ability of the peptide nucleic acid probes to react with target sequences of mycobacterial rRNA was tested by dot blot carried out with rRNA from M. bovis BCG, M. avium and E.coli.

M. bovis BCG (Statens Serum Institut, Denmark) and M. intracellulare (kindly provided by Statens Serum Institut) were grown in Dubos broth (Statens Serum Institut) or on Löwenstein-Jensen slants (Statens Serum Institut) at 37 °C. RNA was isolated from the bacterial cells using TRI-reagent (Sigma) following manufacture's directions. E. coli rRNA was purchased from Boehringer Mannheim, Germany.

25

35

200 ng M. bovis RNA, M. intracellulare RNA and E. coli rRNA were dotted onto membranes (Schleicher & Schüel, NY 13 N), and the membranes were dried and fixed under UV light for 2 minutes.

30 Protocol for dot blot assay

Each of the probes (70 nM probe in hybridisation solution (50 mM Tris, 10 mM NaCl, 10% (w/v) Dextran sulphate, 50% (v/v) glycerol, 5 mM EDTA, 0.1% (w/v) sodium pyrophosphate, 0.2% (w/v) polyvinylpyrrolidone, 0.2% (w/v) Ficoll, pH 7.6.)) were spotted onto a membrane. Hybridisation was continued for 1.5 hours at 55 or 65 °C, respectively. The membranes were rinsed 2 times for 15 minutes in 2 × SSPE buffer (1 x SSPE: 0.15 M NaCl, 10 mM sodium phosphate, 1 mM EDTA, pH 7.4) containing 0.1% SDS at ambient temperature, and subsequently 2 times for 15 minutes in 0.1 × SSPE buffer containing 0.1% SDS at 55 or 65 °C (see Table 1). The membrane was blocked with 0.5% (w/v) casein dissolved in 0.5M NaCl, 0.05M Tris/HCl pH 9.0. Thereafter, the membranes were incubated for 1 hour with rabbit-anti

FITC antibody labelled with alkaline phosphatase (AP) (DAKO K0046 vial A) diluted 1:2000 in 0.5% casein dissolved in 0.5M NaCl, 0.05M Tris/HCl pH 9.0. After incubation, the membranes were washed 3 times 5 minutes with TST buffer (0.05M Tris, 0.5M NaCl, 0.5% (w/v) Tween 20°, pH 9) at ambient temperature. Bound probes were visualised following standard procedures using BClP/NBT, and the visualisation was stopped by incubation for 10 minutes with 10 mM EDTA. The blot was dried at 50 °C.

The results are given in Table 1 below.

TABLE 1

10

	E. coli rRNA		M. bovis BCG RNA		M. intracellulare RNA	
Probe	55 °C	65 °C	55 °C	65 °C	55 °C	65 °C
OK 305	negative	negative	positive	positive	negative	weak
OK 307	negative	negative	positive	positive	negative	weak
OK 309	negative	negative	positive	positive	negative	weak
OK 223	negative	negative	positive	positive	nd	nd
OK 310	negative	negative	negative	positive	negative	negative

nd: Not determined

The results indicate that all five peptide nucleic acid probes are capable of hybridising to target sequence of M. bovis BCG rRNA (as a representative of the MTC group), whereas no hybridisation to E. coli rRNA (as a representative of organisms other than mycobacteria) and no detectable hybridisation to M. intracellulare rRNA were observed (as a representative of the MOTT group).

EXAMPLE 5

20

15

This example illustrates the ability of the peptide nucleic acid probes to penetrate the mycobacterial cell wall and subsequently hybridise to target sequence of mycobacteria of the MTC group and not mycobacteria of the MOTT group, in particular not mycobacteria of the MAC group, or Neisseria gonorrhoeae, by fluorescence *in situ* hybridisation (FISH).

25

Preparation of bacterial slides

M. bovis BCG (Statens Seruminstitut, Denmark), M. avium (kindly provided by Statens Seruminstitut, Denmark), and M. intracellulare (kindly provided by Statens Seruminstitut,

Denmark) were grown in Dubos broth (Statens Seruminstitut, Denmark) or on Löwenstein-Jensen slants (Statens Seruminstitut, Denmark) at 37 °C. N. gonorrhoeae (Statens Seruminstitut, Denmark) was grown on chocolate agar (Statens Seruminstitut, Denmark) at 37 °C with additional 5% CO₂.

5

10

Cultures were smeared onto microscope slides and fixed according to standard procedures. Prior to the hybridisation, the smears were immersed into 80% ethanol for 15 minutes, and subsequently rinsed with water and air dried. This step is not essential for the following hybridisation step, but it is anticipated that it will kill any viable mycobacteria on the slides, and may further serve as an additional fixation step.

Protocol for fluorescence in situ hybridisation (FISH)

- The bacterial slide was covered with a hybridisation solution containing the probe in question.
- 15 2. The slide was incubated in a humid incubation chamber at 45°C or 55°C for 90 minutes.
 - The slide was washed 25 minutes at 45°C or 55°C in prewarmed wash solution (5 mM
 Tris, 145 mM NaCl, pH 10) followed by 30 seconds in water.
- The slide was dried and mounted with IMAGEN Mounting Fluid (DAKO, Copenhagen,
 Denmark)

The hybridisation solution contains 50 mM Tris, 10 mM NaCl, 10% (w/v) Dextran sulphate, 30% (v/v) formamide, 0.1% (v/v) Triton X-100[®], 5 mM EDTA, 0.1% (w/v) sodium pyrophosphate, 0.2% (w/v) polyvinylpyrrolidone, 0.2% (w/v) Ficoll, pH 7.6.

25

30

35

Whenever possible, the applied equipment was heat-treated, and solutions were exposed to 1µl/ml diethylpyrocarbonate (Sigma Chemical Co.) in order to inactivate nucleases.

Microscopically examinations were conducted using a fluorescence microscope (Leica, Wetzlar, Germany) equipped with a 100×/1.20 water objective, a HBO 100 W lamp and a FITC filter set. Mycobacteria were identified as fluorescent, 1 - 10 μm slender, rod-shaped bacilli.

Fluorescein-labelled peptide nucleic acid probes targeting 23S rRNA of the mycobacteria of the MTC group (OK 306, OK 309, OK 446, OK 449) and 16S rRNA of the mycobacteria of the MTC group (OK 310) were tested. Individual probe concentrations and incubation temperatures are listed together with the results in Table 2 and 3.

TABLE 2

	OK 306	OK 309	OK 446	OK 449	
	250nM	250nM	500nM	500nM	
	45°C	45°C	55°C	55°C	
M. bovis BCG	positive	positive	positive	positive	
M. avium	negative	negative	negative	negative	
M. intracellulare	negative	negative	not determined	not determined	
N. gonorrhoeae	negative	negative	not determined	not determined	

TABLE 3

	OK 447	OK 310	OK 306/OK 310
	1μΜ	250nM	500/500nM
	55°C	45°C	55°C
M. bovis BCG	positive	positive	positive
M. avium	negative	negative	negative
M. intracellulare	not determined	negative	negative
N. gonormoeae	not determined	negative	not determined

It can be concluded that the probes are able to penetrate the mycobacterial cell wall of mycobacterium cultures and subsequently hybridise to target rRNA sequence. This makes possible the development of fluorescence in situ hybridisation (FISH) protocols for specific detection of mycobacteria.

10 EXAMPLE 6

15

Test of probes on clinical smears of sputum

The ability of the peptide nucleic acid to penetrate the cell wall of mycobacteria of the MTC group in clinical samples was tested on smears of sputum from suspected cases of tuberculosis (kindly provided by Division of Microbiology, Ramathibodi Hospital, Bangkok, Thailand) by fluorescence in situ hybridisation (FISH). Smears from the same patient were initially evaluated positive by Ziehl-Neelsen staining, which shows only the presence of acid fast bacilli, not whether these are mycobacteria of the MTC group.

Fluorescein-labelled peptide nucleic acid probes targeting 23S rRNA of the mycobacteria of the MTC group (OK 306, OK 446, OK 449) and 16S rRNA of the mycobacteria of the MTC group (OK 310) were used. Furthermore, a random peptide nucleic acid probe (a 15-mer wherein each position may be A, T, C or G (obtained from Millipore Corporation, Bedford, MA, USA) was added to the hybridisation solution in order to increase the signal-to-noise ratio.

FISH was carried out at 55 °C as described in Example 5. Applied probe concentrations are listed together with the results in Table 4 and 5.

TABLE 4

Sample	OK 446/Random	OK 449/Random	Ziehl-Neelsen
number	1μМ/50μМ	1μΜ/50μΜ	staining
285	Positive	Positive	4+
335	Positive	Eq.	2+
345	Positive	Positive	3+
224	Positive	Positive	3+
297	Negative	Eq.	2+
179	Negative	Negative	4+
247	Negative	Negative	2+
255	Positive	Positive	2+
202	Eq.	Positive	2+

5

TABLE 5

Sample number	OK 306/OK 310 500/500 nM	Ziehl-Neelsen staining		
213	Positive	4+		
292	Positive	4+		
159	Positive	3+		
287	Positive	3+		

Smears stained by Ziehl-Neelsen staining were examined with a 100x objective and scored according to the following method: -: 0 bacilli, +/-: 1-200 per 300 fields, 2+: 1-9 per 10 fields, 3+: 1-9 per field, 4+: >9 per field.

Positive: Several mycobacteria were identified in the smear. Negative: No fluorescent mycobacteria were identified in the smear. Eq. Few (1-3) fluorescent mycobacteria were identified in the smear.

It appears from the table that the peptide nucleic acid probes are able to penetrate and subsequently hybridise to target sequence of mycobacteria of the MTC-group in AFB-positive sputum smears. The fact that not all AFB-positive sputum smears are found positive with applied probes indicate that not all AFB-positive sputum smears contains mycobacteria of the MTC-group.

EXAMPLE 7

20

15

The reactivity and specificity of selected peptide nucleic acid probes for detecting

mycobacteria of the MTC group as well as probes for detecting mycobacteria of the MOTT group were evaluated by fluorescence in situ hybridisation (FISH) on control smears prepared from cultures of different mycobacterium species. The mycobacterium species were selected so as to be representative for the mycobacterium genus as well as to include clinically relevant species.

M. tuberculosis (ATCC 25177), M. bovis BCG (ATCC 35734), M. intracellulare (ATCC 13950), M. avium (ATCC 25292), M. kansasii (ATCC12479), M. gordonae (ATCC 14470), M. scrofulaceum (ATCC 19981), M. abscessus (ATCC19977), M. marinum (ATCC 927), M. simiae (ATCC 25575), M. szulgai (ATCC 35799), M. flavescens (ATCC 23033), M. fortuitum (ATCC 43266) and M. xenopi (ATCC19250) were grown at Dubos broth (Statens Serum Institut) at 37 °C with the exception of M. marinum which was grown at 32 °C.

Smears were prepared as described in Example 5. FISH was carried out as described below.

15

10

5

Protocol for fluorescence in situ hybridisation (FISH)

- The bacterial slide was covered with a hybridisation solution containing the probe in question.
- The slide was incubated in a humid incubation chamber at 55°C for 90 minutes.
- 20 3. The slide was washed 30 minutes at 55°C in prewarmed wash solution (5 mM Tris, 15 mM NaCl, 0.1% (v/v), Triton X-100[®], pH 10) followed by 30 seconds in water.
 - The slide was dried and mounted with IMAGEN Mounting Fluid (DAKO, Copenhagen, Denmark)
- The hybridisation solution contained 50 mM Tris, 10 mM NaCl, 10% (w/v) Dextran sulphate, 30% (v/v) formamide, 0.1% (v/v) Triton X-100°, 5 mM EDTA, 0.1% (w/v) sodium pyrophosphate, 0.2% (w/v) polyvinylpyrrolidone, and 0.2% (w/v) Ficoll, pH 7.6. To avoid non-specific binding of the labelled peptide nucleic acid probe, 1-5 μM of non-labelled, non-sense peptide nucleic acid probe was added to the hybridisation solution (OK 507/modified Seq ID no 121 and/or OK 714/modified Seq ID no 122).

Whenever possible, the applied equipment was heat-treated, and solutions were exposed to 1µl/ml diethylpyrocarbonate (Sigma Chemical Co.) in order to inactivate nucleases.

Microscopic examinations were conducted using a fluorescence microscope (Leica, Wetzlar, Germany) equipped with a 100×/1.30 oil objective, a HBO 100 W lamp and a FITC/TRITC dual band filter set. Mycobacteria were identified on basis of both fluorescence (strong, medium, weak, no) and morphology (1-10 μm slender, rod-shaped bacilli. Mycobacteria of the MOTT

group may appear pleomorphic, ranging in appearance from long rods to coccoid forms)

Probe concentrations are listed together with the results in Table 6 and 7 (probes targeting mycobacteria of the MTC group) and Table 8 (probes targeting to mycobacteria of the MOTT group).

TABLE 6

	OK 450	OK 682	OK 689	OK 688	OK 660
	25 nM	100 nM	100 nM	250 nM	100 nM
M. tuberculosis	+++	+++	+++	+++	+++
M. bovis BCG	+++	+++	+++	+++	+++
M. intracellulare	-	-	-	•	-
M. avium	-	•	-	•	-
M. kansasii	++	•	•	•	_
M. gordonae	-	-	-	-	-
M. scrofulaceum	+++	•	-	•	-
M. abscessus	-	-	•		+
M. marinum	+++	-	+	+	+++
M. simiae	-	-	-	•	-
M. szulgai	+++	:-	-		· -
M. flavescens	 	++	-	-	
M. fortuitum	1 -	+	-	-	
M. xenopi	 - 	++			

⁺⁺⁺ strong fluorescence, ++ medium fluorescence, + weak fluorescence, - no fluorescence

TABLE 7

Mycobacteria	OK 655	OK 448	OK 654	OK 446
	150 nM	50 nM	100 nM	25 nM
M. tuberculosis	+++	+++	+++	+++
M. bovis BCG	+++	+++	+++	+++
M. intracellulare	1 -	-	•	•
M. avium	•	•		-
M. kansasii	-	* -	•	-
M. gordonae	-	•	-	-
M. scrofulaceum	-	-	-	-
M. abscessus	-	-	+	
M. marinum	- 1	•	+	+++
M. simiae	-	-	-	•
M. szulgai	1 - 1	_	•	-
M. flavescens	•	-	-	•
M. fortuitum	-	-	-	-
M. xenopi	- 1	-	-	-

⁺⁺⁺ strong fluorescence, ++ medium fluorescence, + weak fluorescence, - no fluorescence

TABLE 8

OK 612	OK 624	OK 623
100 nM	100 nM	100 nM
•	-	+
-	-	•
	++	++
+++	+++	+++
-	-	+++
-	++	++ -
-	++	++
-	++	+++
	•	•
-	++	+++
- !	-	+++
-	•	•
-	++	•
-	-	•
	-	100 nM 100 nM ++ +++ +++ ++ - ++ - ++ - +

⁺⁺⁺ strong fluorescence, ++ medium fluorescence, + weak fluorescence, - no fluorescence

Each of probes indicated in Table 6, 7 and 8 was further investigated with regard to hybridisation to other common respiratory bacteria, namely Corynebacterium spp.,

Fusobacterium nucleatum, Haemophilus influenzae, Klebsiella pneumoniae, Pseudomonas aeruginosa, Propionibacterium acnes, Streptococcuc pneumoniae, Staphylococcus aureus, Brahamella catarrahalis, Escherichia coli, Neisseria spp., Actinobacter calcoaceticus, Actinomyces spp., Enterobacter aerogenes, Proteus mirabilis, Pseudomonas maltophilia, Streptocussuc viridans, and Norcardia asteroides. No cross-hybridisation was observed by fluorescence in situ hybridisation to any of these bacteria in the case of OK 682, OK 654, OK 655, OK 688, OK 660, OK 612, OK 624 and OK 623. Some cross-reactivity was observed in the case of OK 446 (to P. acnes), OK 448 (to P. acnes and B. catarrhalis), and OK 450 (to P. acnes and B. catarrhalis).

10

15

20

35

5

Table 6 and 7 shows that none of the MTC probes cross-react with M. intracellulare and/or M. avium, but indeed strongly with M. tuberculosis and M. bovis BCG. As shown in Table 8, both OK 624 and OK 623 hybridise to M. intracellulare and M. avium which are both members of the MAC group, whereas none of them hybridise to M. tuberculosis or M. bovis BCG. OK 612 hybridises to M. avium only. It should be noted that the aligned sequence of M. intracellulare has just one nucleobase difference to the target sequence of M. avium, see Figure 4K.

The data support the use of the methodology described in claim 3 and 4 and exemplified in Example 2 for design of peptide nucleic acid probes that are capable of hybridising to target sequence of one or more mycobacterium species and not to other mycobacterium species having at least one nucleobase difference to the target sequence.

EXAMPLE 8

To study the usefulness of the peptide nucleic acid probes in distinguishing between mycobacteria of the MTC group and mycobacteria of the MOTT group, the probes were tested on smears of mycobacterium-positive cultures prepared from 34 + 28 clinical samples (sputum samples, other respiratory samples and extrapulmonary samples) from individuals suspected of tuberculosis or other mycobacterial infections (kindly provided by the Mycobacterium Department, Statens Serum Institut, Denmark). Complex/species identification data obtained with the AccuProbe tests from Gen-Probe Inc., USA were available for each sample.

Table 9 shows the results obtained with four different peptide nucleic acid probes targeting mycobacteria of the MTC group (OK 682, OK 660, OK 688 and OK 689) and one probe targeting mycobacteria of the MOTT group (OK 623), and Table 10 shows the results obtained with two peptide nucleic acid probes targeting mycobacteria of the MOTT group (OK 623 and OK 612) and a mixture of two probes targeting mycobacteria of the MTC group (OK 688 and OK 689). Data are arranged according to the results obtained by AccuProbe. Sample

preparation, hybridisation and visualisation were performed as described in Example 7.

TABLE 9

Complex/	OK 623	OK 682	OK 660	OK 688	OK 689
species (n)	25 nM	100 nM	100 nM	250 nM	100 nM
	n _p	n _p	n _p	n _p .	n _p
MTC (23)	0	23	23	23	23
M. avium (5)	5	0	0	0	0
M. gordonae (3)	3	0	0	0	0
Unknown (3)	3	0	0	0	0

n_p denotes number of positive samples.

The term "unknown" means that the sample not contains mycobacteria of the MTC group, or mycobacteria of the MAC group according the AccuProbe test, but further species identification was not performed.

TABLE 10

10

Complex/	OK 623	OK 612	OK 688/OK 689
species (n)	25nM	100 nM	50 nM/50 nM
	n _p	n _p	n _p
MTC (17)	0		16
M. avium (2)	2	2	0
M. gordonae (4)	3	0	0
Unknown (5)	5	0	0

n_p denotes number of positive samples.

The term "unknown" means that the sample not contains mycobacteria of the MTC group, or mycobacteria of the MAC group according to the AccuProbe test, but further species identification was not performed.

- The results shown in Table 9 are in conformity with the complex/species identification performed with the AccuProbe tests, and thus confirm that peptide nucleic acid probes can be used to determine whether an infection is caused by mycobacteria of the MTC group or by mycobacteria of the MOTT group.
- From the results in Table 10, it can be seen that it is possible to differentiate between mycobacteria of the MTC group and mycobacteria of the MOTT group with 100% specificity and 91-94% sensitivity relative to results obtained by the AccuProbe tests. Furthermore, OK 612 is very suitable for specific identification of M. avium among those being positive for mycobacteria of the MOTT group as the result is positive in the case of M. avium and negative in the other cases of mycobacteria of the MOTT group.

EXAMPLE 9

Direct detection of mycobacteria in clinical smears of sputum

This example demonstrates the ability of the peptide nucleic acid to detect and identify mycobacteria directly in AFB-positive sputum samples from suspected cases of tuberculosis (kindly provided by Division of Microbiology, Ramathibodi Hospital, Bangkok, Thailand) and suspected cases of other mycobacterial infections (kindly provided by Clinical Microbiology Dept., Rigshospitalet, Copenhagen, Denmark) by FISH is shown.

The clinical smears were prepared according to the procedure described in Example 5, and FISH was performed as described in Example 7. The results are shown in Table 11.

TABLE 11

	OK 623	OK 654	OK 655	OK 682	OK 688	OK 689
Sample no.	25 nM	100 nM	150 nM	100 nM	250 nM	100 nM
1	•	++	++	++	++	++
175	•	++	nd	nd	++	++
459	-	•	nd	nd	-	-
166	•	•	•	nd	•	•
268	-	++	++	++	++	++
34267	++	•	-	-	-	-

nd: not determined

15

20

25

30

+++ strong fluorescence, ++ medium fluorescence, + weak fluorescence, - no fluorescence

It appears from examples in Table 11 that AFB-positive sputum smears were evaluated positive for mycobacteria of the MTC group (sample numbers 1, 175, and 268), positive for mycobacteria of the MOTT group (sample number 37267), or negative for mycobacteria (sample numbers. 459 and 166) by the applied probes. Thus, PNA-probes are useful reagents for specific identification of mycobacteria directly in sputum smears by fluorescence in situ hybridisation. AFB-positive sputum samples that are negative with all probes may be explained in three ways: a) the sample may contain mycobacteria not detected by the probes, e.g. M. fortuitum, b) the sample may contain other acid-fast bacteria than mycobacteria, or c) the mycobacteria in the sample lack or have a strongly reduced content of rRNA due to for example antibiotic treatment.

In conclusion, direct identification of mycobacteria in smear-positive sputum samples by peptide nucleic acid-based fluorescence in situ hybridisation combines simplicity and morphological advantages of current staining methods with concominant species identification, and will thus allow clinical microbiology laboratories to benefit from the

advantages offered by molecular techniques to provide crucial information pertaining to therapy and patient management.

EXAMPLE 10

5

This example demonstrates simultaneous detection and identification of mycobacteria of the MTC group and mycobacteria of the MOTT group using differently labelled probes targeting mycobacteria of the MTC group and mycobacteria of the MOTT group, respectively, by fluorescence in situ hybridisation.

10

Control smears of different mycobacterium species were prepared as described in Example 5. In addition, smears containing a mixture of M. tuberculosis and M. avium were prepared (Table 8, last row). FISH was performed as described in Example 7.

A rhodamine-labelled peptide nucleic acid probe targeting 16S rRNA of mycobacteria of the MTC group (OK 702) and a fluorescein-labelled peptide nucleic acid probe targeting 16S rRNA of mycobacteria of the MOTT group (OK 623) were applied simultaneously in the concentrations listed in Table 12 together with the results.

20 TABLE 12

25

30

Mycobacterium species	OK 623/OK 702	
	25/250 nM	
M. tuberculosis	- (G)/ +++ (R)	
M. bovis BCG	- (G)/ +++ (R)	
M. avium	+++ (G)/ - (R)	
M. intracellulare	+++ (G)/ - (R)	
M. kansasii	+++ (G)/ - (R)	
M. avium / M. tuberculosis	+++ (G)/+++ (R)	

⁺⁺⁺ strong fluorescence - no fluorescence

Mycobacteria of the MTC group, i.e. M. tuberculosis and M. bovis, were observed as green fluorescent mycobacteria, whereas mycobacteria of the MOTT group, i.e. M. avium, M. intracellulare and M. kansasii, were observed as red fluorescent mycobacteria. Mycobacteria in the M. avium/M. tuberculosis mixture were identified by a mixture of both green fluorescent mycobacteria and red fluorescent mycobacteria.

The results show that it is possible to distinguish between different Mycobacterium species in

G green fluorescence, R red fluorescence

one smear using a mixture of differently labelled probes. Such simultaneous detection and identification of mycobacteria may further be extended to comprise three or more differently labelled peptide nucleic acid probes.

5 EXAMPLE 11

The ability of a peptide nucleic acid probes to hybridise to precursor rRNA and further to distinguish between precursor rRNA of M. tuberculosis and precursor rRNA of M. avium was investigated by fluorescence in situ hybridisation.

10

20

30

Smears were prepared as described in Example 5 and FISH were carried out as described in Example 7 using a fluorescein-labelled probe targeting precursor rRNA of M. tuberculosis (OK 749). The results are given in Table 13.

15 TABLE 13

1000 nM
+
-

⁺ weak fluorescence - no fluorescence

From the results, it can be concluded that it is possible to detect precursor rRNA, and further that is possible to distinguish between precursor rRNA from different mycobacterium species. The application of peptide nucleic acid targeting precursor rRNA may be particularly useful for measuring the mycobacterial growth and thus be an indicator of the viability of the mycobacteria. This would in particular be important for monitoring of the effect of antibiotics in relation to both treatment of tuberculosis and drug susceptibility studies.

25 EXAMPLE 12

The ability of peptide nucleic acid probes for differentiation of drug susceptible and drug resistant mycobacteria was evaluated using a fluorescein-labelled probe targeting the wild type sequence of 23S rRNA of M. avium and M. intracellulare together with rhodamine-labelled probes targeting single point mutations associated with macrolide resistance in M. avium and M. intracellulare.

Smears were prepared as described in Example 5 from cultures of M. avium (ATCC no. 25292) and M. intracellulare (ATCC no. 13950). These strains are anticipated to contain the

wild type sequence of rRNA. Macrolide resistant variants were not available. FISH was carried out as described in Example 7 using a fluorescein-labelled peptide nucleic acid probe targeting wild type 23S rRNA (OK 745) and a mixture of rhodamine-labelled peptide nucleic acid probes targeting the three possible mutations at position 2568 (OK 746) and at position 2569 (OK 747) of M. avium 23S rDNA of GenBank entry X52917 (see Figure 6). The results are given in Table 14.

TABLE 14

10

15

20

25

Mycobacterium species	OK 745/OK 746/OK 747
	500/500/500 nM
M. avium (wild type)	+++ (G)/ - (R)
M. intracellulare (wild type)	+++ (G)/ - (R)

⁺⁺⁺ strong fluorescence - no fluorescence

G green fluorescence, R red fluorescence

OK 748 and OK 747 are each a mixture of three single point mutation probes

The results in Table 14 show that M. avium and M. intracellulare are detected with the fluorescein-labelled probe (OK 745) targeting M. avium and M. intracellulare wild types and not detected with the mixture of rhodamine-labelled probes (OK 746 and OK 747) targeting single point mutations associated with macrolide resistance. Such peptide nucleic acid probes targeting the wild type and drug resistant variants, respectively, may be important tools for both the prediction of an efficient therapy as well as for monitoring the effect of the treatment.

EXAMPLE 13

To illustrate the speed with which peptide nucleic acid probes penetrate the mycobacterial cell wall and subsequently hybridise to their target sequence the protocol described in Example 7 was modified to 15 minutes hybridisation time and the results compared with 90 minutes hybridisation time. Smears were prepared as described in Example 5. The results are given in Table 15.

TABLE 15

	ОК	623	ОК	689
	25 nM		100 nM	
	15 min	90 min	15 min	90 min
M. tuberculosis			++	+++
M. avium	++	+++		

⁺⁺⁺ strong fluorescence ++ medium fluorescence

The data presented in Table 15 show that hybridisation by peptide nucleic acid probes inside the mycobacterial cells is accomplished in a very short time resulting in a detectable signal after just 15 minutes incubation. Thus, the use peptide nucleic acid probes makes possible the development of very fast fluorescence in situ hybridisation protocols.

10 EXAMPLE 14

To describe the ability of very short peptide nucleic acid probes to hybridise to target sequences, a 12-mer peptide nucleic acid probe labelled with fluorescein (OK 575) was tested by fluorescence in situ hybridisation (FISH).

15

25

Smears were prepared as described in Example 5 and FISH were carried out as described in Example 7. The results are given in Table 16.

TABLE 16

Mycobacterium	OK 575
	50 nM
M. tuberculosis	+
M. bovis BCG	++
M. avium	-
M. intracellulare	*
M. kansasii	-

20 ++ medium fluorescence + weak fluorescence - no fluorescence

The results in table 17 shows that a 12-mer peptide nucleic acid probe is capable of hybridising specifically to target sequences under the same stringency conditions as 15-mers. A lower florescence intensity is obtained as the T_m for a 12-mer peptide nucleic acid probe is lower than T_m for a 15-mer peptide nucleic acid probe.

⁺ weak fluorescence - no fluorescence

WO 98/15648 PCT/DK97/00425 54

The data clearly suggest that by lowering the stringency condition, e.g. by decreasing the hybridisation/washing temperature and/or the concentration of formamide, even shorter probes may be applied for detection of mycobacteria provided that specific sequences of such can be designed.

CLAIMS

10

15

20

25

- 1. Peptide nucleic acid probe for detecting a target sequence of one or more mycobacteria optionally present in a sample, said probe being capable of hybridising to a target sequence of mycobacterial rDNA, precursor rRNA or rRNA forming detectable hybrids, and a mixture of such probes.
- 2. Peptide nucleic acid probe according to claim 1, said probe being capable of hybridising to a target sequence of mycobacterial rDNA, precursor rRNA, or 23S, 16S or 5S rRNA forming detectable hybrids, and a mixture of such probes.
- 3. Peptide nucleic acid probe according to claim 1 or 2, said probe being capable of hybridising to a target sequence of mycobacterial rDNA, precursor rRNA, or 23S, 16S or 5S rRNA forming detectable hybrids, said target sequence being obtainable by
- (a) comparing the nucleobase sequences of said mycobacterial rRNA or rDNA of one or more mycobacteria to be detected with the corresponding nucleobase sequence of organism(s), in particular other mycobacteria, in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished,
- (b) selecting a target sequence of said rRNA or rDNA which includes at least one nucleobase differing from the corresponding nucleobase of the organism(s), in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished, and
- (c) determining the capability of said probe to hybridise to the selected target sequence to form detectable hybrids, and a mixture of such probes.
- 4. Peptide nucleic acid probe according to claim 1 or 2, said probe being capable of hybridising to a target sequence of mycobacterial rDNA, precursor rRNA or 23S, 16S or 5S rRNA forming detectable hybrids, said probe being obtainable by
- (a) comparing the nucleobase sequences of said mycobacterial rRNA or rDNA of one or more
 mycobacteria to be detected with the corresponding nucleobase sequence of organism(s), in particular other mycobacteria, in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished.

WO 98/15648 PCT/DK97/00425 56

- (b) selecting a target sequence of said rRNA or rDNA which includes at least one nucleobase differing from the corresponding nucleobase of the organism(s), in particular other mycobacteria, from which said one or more mycobacteria are to be distinguished.
- 5 (c) synthesising said probe, and

10

30

35

(d) determining the capability of said probe to hybridise to the selected target sequence to form detectable hybrids, and a mixture of such probes.

5. Peptide nucleic acid probe according to any one of claims 1 to 4 for detecting a target sequence of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) or for detecting a target sequence of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT) optionally present in a sample, which probe comprises from 6 to 30 polymerised peptide nucleic acid moieties, said probe being capable of

hybridising to a target sequence of mycobacterial rDNA, precursor rRNA or 23S, 16S or 5S rRNA forming detectable hybrids.

and a mixture of such probes.

20 6. Peptide nucleic acid probe according to any one of claims 1 to 5 for detecting a target sequence of rDNA, precursor rRNA or 23S, 16S or 5S rRNA of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) or for detecting a target sequence of rDNA, precursor rRNA or 23S, 16S or 5S rRNA of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT) optionally present in a 25 sample, which probe comprises from 10 to 30 polymerised moieties of formula (I)

wherein each X and Y independently designate O or S,

each Z independently designates O, S, NR¹, or C(R¹)₂, wherein each R¹ independently designate H, C₁₋₆ alkyl, C₁₋₆ alkenyl, C₁₋₆ alkynyl,

each R², R³ and R⁴ designate independently H, the side chain of a naturally occurring amino acid, the side chain of a non-naturally occurring amino acid, C14 alkyl, C14 alkenyl or C14 alkynyl, or a functional group, each Q independently designates a naturally occurring nucleobase, a non-naturally occurring nucleobase, an intercalator, a nucleobase-binding

5

10

15

group, a label or H,

with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with the target sequence of said mycobacterial rDNA, precursor rRNA or 23S, 16S or 5S rRNA,

and a mixture of such probes.

7. Peptide nucleic acid probe according to any one of claims 1 to 6 for detecting a target sequence of 23S rRNA of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) optionally present in a sample, which probe comprises from 10 to 30 polymerised moieties of formula (I) as defined in claim 6,

with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase that is complementary to a nucleobase of M. tuberculosis 23S rRNA differing from the corresponding nucleobase of at least M. avium located within the following domains

Positions 149-158 in Figure 1A,

Positions 220-221 in Figure 1A,

20 Positions 328-361 in Figure 1A and Figure 1B.

Positions 453-455 in Figure 1B,

Positions 490-501 in Figure 1B,

Positions 637-660 in Figure 1C,

Positions 706-712 in Figure 1D,

25 Positions 762-789 in Figure 1D,

Position 989 in Figure 1D,

Positions 1068-1072 in Figure 1D,

Position 1148 in Figure 1E,

Positions 1311-1329 in Figure 1E,

30 Positions 1361-1364 in Figure 1F,

Position 1418 in Figure 1F,

Positions 1563-1570 in Figure 1F,

Positions 1627-1638 in Figure 1G,

Positions 1675-1677 in Figure 1G,

35 Position 1718 in Figure 1G,

Positions 1734-1740 in Figure 1H,

Positions 1967-1976 in Figure 1H,

Positions 2403-2420 in Figure 1H,

Positions 2457-2488 in Figure 1I,
Positions 2952-2956 in Figure 1I,
Positions 2966-2969 in Figure 1J,
Positions 3000-3003 in Figure 1J or
Positions 3097-3106 in Figure 1J.

and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with a target sequence of said mycobacterial 23S rRNA, and a mixture of such probes.

10

8. Peptide nucleic acid probe according to any one of claims 1 to 6 for detecting a target sequence of 16S rRNA of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) optionally present in a sample, which probe comprises from 10 to 30 polymerised moieties of formula (I) as defined in claim 6,

15

with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase that is complementary to a nucleobase of M. tuberculosis 16S rRNA differing from the corresponding nucleobase of at least M. avium located within the following domains

20

35

Positions 76-79 in Figure 2A,
Positions 98-101 in Figure 2A,
Positions 135-136 in Figure 2 A,
Positions 194-201 in Figure 2B,
Positions 222-229 in Figure 2B,
Position 242 in Figure 2B,
Position 474 in Figure 2C,
Positions 1136-1145 in Figure 2C,
Positions 1271-1272 in Figure 2C,
Positions 1287-1292 in Figure 2D,
Position 1313 in Figure 2D, or
Position 1334 in Figure 2D,

and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with a target sequence of said mycobacterial 16S rRNA, and a mixture of such probes.

9. Peptide nucleic acid probe according to any one of claims 1 to 6 for detecting a target

WO 98/15648 PCT/DK97/00425

sequence of 5S rRNA of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) optionally present in a sample, which probe comprises from 10 to 30 polymerised moieties of formula (I) as defined in claim 6,

- with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase that is complementary to a nucleobase of M. tuberculosis 5S rRNA differing from the corresponding nucleobase of at least M. avium located within the following domain
- 10 Positions 86-90 in Figure 3

and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with a target sequence of said mycobacterial 5S rRNA, and a mixture of such probes.

15

10. Peptide nucleic acid probe according to any one of claims 1 to 8 for detecting a target sequence of 23S or 16S rRNA of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) optionally present in a sample, which probe comprises from 10 to 30 polymerised moieties of formula (I) as defined in claim 6,

20

with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase that is complementary to a nucleobase of M. tuberculosis 23S or 16 S rRNA differing from the corresponding nucleobase of at least M. avium located within the following domains

25

Positions 149-158 in Figure 1A,
Positions 328-361 in Figure 1A and Figure 1B,
Positions 490-501 in Figure 1B,
Positions 637-660 in Figure 1C,

Positions 762-789 in Figure 1D,
Positions 1068-1072 in Figure 1D,
Positions 1311-1329 in Figure 1E,
Positions 1361-1364 in Figure 1F,
Positions 1563-1570 in Figure 1F,
Positions 1627-1638 in Figure 1G,
Positions 1734-1740 in Figure 1H,
Positions 2457-2488 in Figure 1I,

Positions 2952-2956 in Figure 11,

Positions 3097-3106 in Figure 1J, Positions 135-136 in Figure 2 A, or Positions 1287-1292 in Figure 2D,

- and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with a target sequence of said mycobacterial 23S or 16S rRNA, and a mixture of such probes.
- 11. Peptide nucleic acid probe according to any one of claims 1 to 6 for detecting a target
 sequence of 23S rRNA of one or more mycobacteria other than mycobacteria of the
 Mycobacterium tuberculosis Complex (MOTT) optionally present in a sample, which probe comprises from 10 to 30 polymerised moieties of formula (I) as defined in claim 6.
- with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of
 which a subsequence includes at least one nucleobase that is complementary to a
 nucleobase of M. avium 23S rRNA differing from the corresponding nucleobase of at least M.
 tuberculosis located within the following domains

Positions 99-101 in Figure 4A,

20 Position 183 in Figure 4A,

Positions 261-271 in Figure 4A,

Positions 281-284 in Figure 4B,

Positions 290-293 in Figure 4B,

Positions 327-335 in Figure 4B,

25 Positions 343-357 in Figure 4B,

Positions 400-405 in Figure 4B and Figure 4C.

Positions 453-462 in Figure 4C,

Positions 587-599 in Figure 4C,

Positions 637-660 in Figure 4D,

30 Positions 704-712 in Figure 4D,

Positions 763-789 in Figure 4E,

Positions 1060-1074 in Figure 4E,

Positions 1177-1185 in Figure 4E,

Positions 1259-1265 in Figure 4F,

35 Positions 1311-1327 in Figure 4F,

Positions 1345-1348 in Figure 4F,

Positions 1361-1364 in Figure 4G,

Positions 1556-1570 in Figure 4G,

Positions 1608-1613 in Figure 4H,

Positions 1626-1638 in Figure 4H,

Positions 1651-1659 in Figure 4H,

Positions 1675-1677 in Figure 4H,

5 Positions 1734-1741 in Figure 4H,

Positions 1847-1853 in Figure 4I,

Positions 1967-1976 in Figure 4I,

Positions 2006-2010 in Figure 4I,

Positions 2025-2027 in Figure 4I,

10 Positions 2131-2132 in Figure 4J,

Positions 2252-2255 in Figure 4J,

Positions 2396-2405 in Figure 4J and Figure 4K,

Positions 2416-2420 in Figure 4K,

Positions 2474-2478 in Figure 4K,

15 Position 2687 in Figure 4K,

Position 2719 in Figure 4K,

Position 2809 in Figure 4L,

Positions 3062-2068 in Figure 4L, or

Positions 3097-3106 in Figure 4L,

20

and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with a target sequence of said mycobacterial 23S rRNA, and a mixture of such probes.

- 12. Peptide nucleic acid probe according to any one of claims 1 to 6 for detecting a target sequence of 16S rRNA of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT) optionally present in a sample, which probe comprises from 10 to 30 polymerised moieties of formula (I) as defined in claim 6,
- with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase that is complementary to a nucleobase of M. avium 16S rRNA differing from the corresponding nucleobase of at least M. tuberculosis located within the following domains
- Positions 135-136 in Figure 5A, Positions 472-475 in Figure 5A, Positions 1136-1144 in Figure 5A, Positions 1287-1292 in Figure 5B,

10

35

Position 1313 in Figure 5B, or Position 1334 in Figure 5B,

and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with a target sequence of said mycobacterial 16S rRNA, and a mixture of such probes.

13. Peptide nucleic acid probe according to any one of claims 1 to 6, 11 and 12 for detecting a target sequence of 23S or 16S rRNA of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT) optionally present in a sample, which probe comprises from 10 to 30 polymerised moleties of formula (I) as defined in claim 6,

with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase that is complementary to a nucleobase of M. avium 23S or 16S rRNA differing from the corresponding nucleobase of at least M. tuberculosis located within the following domains

Positions 99-101 in Figure 4A, Positions 290-293 in Figure 4B,

20 Positions 400-405 in Figure 4B and Figure 4C.

Positions 453-462 in Figure 4C.

Positions 637-660 in Figure 4D.

Positions 763-789 in Figure 4E,

Positions 1311-1327 in Figure 4F,

25 Positions 1361-1364 in Figure 4G,

Positions 1734-1741 in Figure 4H.

Positions 2025-2027 in Figure 4I.

Positions 2474-2478 in Figure 4K,

Positions 3062-2068 in Figure 4L, or

30 Positions 1287-1292 in Figure 5B,

and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with a target sequence of said mycobacterial 23S or 16S rRNA, and a mixture of such probes.

14. Peptide nucleic acid probe according to any one of claims 1 to 6 for detecting a target sequence of 23S, 16S or 5S rRNA of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC) or for detecting a target sequence of 23S, 16S or 5S rRNA of

one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT) optionally present in a sample, which probe comprises from 10 to 30 polymerised moleties of formula (I) as defined in claim 6,

- with the proviso that the Qs of adjacent moleties are selected so as to form a sequence of which a subsequence includes at least one nucleobase that is complementary to a nucleobase that differs from the corresponding nucleobase of 23S, 16S or 5S rRNA of said one or more mycobacteria located within the following domains
- 10 positions 2568-2569 in Figure 6,

15

25

30

35

Position 452 in Figure 7, Positions 473-477 in Figure 7, or Positions 865-866 in Figure 7,

and further with the proviso that the probe comprising such subsequence is capable of forming detectable hybrids with the target sequence of said mycobacterial 23S, 16S or 5S rRNA, and a mixture of such probes.

20 15. Peptide nucleic acid probe according to any one of claims 1 to 14 of formula (II), (III), or (IV)

$$\begin{array}{c|c} & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & \\ & & & \\ & \\ & & \\ & & \\ & & \\ & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ & & \\ &$$

$$\sum_{\mathbf{Z}} N \bigvee_{\mathbf{R^4}} O$$

wherein Z, R², R³, and R⁴, and Q is as defined in claim 6 with the provisos defined in claims 6

to 14,

and a mixture of such probes.

16. Peptide nucleic acid probe according to any one of claims 1 to 15, wherein Z is NH, NCH, or O, each R², R³ and R⁴ independently designate H or the side chain of a naturally occurring amino acid, the side chain of a non-naturally occurring amino acid, or C14 alkyl, and each Q is a naturally occurring nucleobase or a non-naturally occurring nucleobase with the provisos defined in claims 6 to 14, and a mixture of such probes.

10

5

- 17. Peptide nucleic acid probe according to any one of claims 1 to 16, wherein Z is NH or O, and R2 is H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, and Q is a nucleobase selected from thymine, adenine, cytosine, guanine, uracil, iso-C and 2,6diaminopurine with the provisos defined in claims 6 to 14,
- and a mixture of such probes.
 - 18. Peptide nucleic acid probe according to any one of claims 1 to 17 of formula (V)

wherein R4 is H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, and Q is as defined in claim 17 with the provisos defined in claims 6 to 14, and a mixture of such probes.

- 25
 - 19. Peptide nucleic acid probe according to any one of claims 1 to 18 further comprising one or more labels and a mixture of such probes, which labels may be mutually identical or different, which probes optionally may comprise one or more linkers, and which probes may be mutually identical or different with the provisos defined in claims 6 to 14.
 - 20. Peptide nucleic acid probe according to any one of claims 1 to 19 for detecting a target sequence of one or more mycobacteria, the nucleobase sequence of said probe being substantially complementary to the nucleobase sequence of said target sequence.

35

30

21. Peptide nucleic acid probe according to any one of claims 1 to 20 for detecting a target sequence of one or more mycobacteria, the nucleobase sequence of said probe being complementary to the nucleobase sequence of said target sequence.

22. Peptide nucleic acid probes according to any one of claims 1 to 21, wherein the Qs of adjacent moieties are selected so as to form the following subsequences

5	AGA TGC GGG TAG CAC (selected from positions 149-158 in Figure 1A),	(Seq ID no 1)
	TGT TTT CTC CTC CTA (selected from positions 220-221 in Figure 1A),	(Seq ID no 2)
	ACT GCC TCT CAG CCG (selected from positions 328-361 in	
	Figure 1A and Figure 1B),	(Seq ID no 3)
	TGA TAC TAG GCA GGT (selected from positions 453-455 in Figure 1B),	(Seq ID no 4)
10	CGG ATT CAC AGC GGA (selected from positions 490-501 in Figure 1B),	(Seq ID no 5)
	TCA CCA CCC TCC TCC (selected from positions 637-660 in Figure 1C),	(Seq ID no 6)
	TTA ACC TTG CGA CAT (selected from positions 706-712 in Figure 1C),	(Seq ID no 7)
	ACT ATT CAC ACG CGC (selected from positions 762-789 in Figure 1D),	(Seq ID no 8)
	CTC CGC GGT GAA CCA (selected from position 989 in Figure 1D),	(Seq ID no 9)
15	GCT TTA CAC CAC GGC (selected from positions 1068-1072 in Figure 1D),	(Seq ID no 10)
	ACG CTT GGG GGC CTT (selected from position 1148 in Figure 1E),	(Seq ID no 11)
	CCA CAC CCA CCA CAA (selected from positions 1311-1329 in Figure 1E),	(Seq ID no 12)
	CCG GTG GCT TCG CTG (selected from positions 1361-1364 in Figure 1F),	(Seq ID no 13)
	ACT TGC CTT GTC GCT (selected from position 1418 in Figure 1F),	(Seq ID no 14)
20	GAT TCG TCA CGG GCG (selected from positions 1563-1570 in Figure 1F),	(Seq ID no 15)
	AAC TCC ACA CCC CCG (selected from positions 1627-1638 in Figure 1G),	(Seq ID no 16)
	ACT CCA CAC CCC CGA (selected from positions 1627-1638 in Figure 1G),	(Seq ID no 17)
	ACC CCT TCG CTT GAC (selected from positions 1675-1677 in Figure 1G),	(Seq ID no 18)
	CTT GCC CCA GTG TTA (selected from position 1718 in Figure 1G),	(Seq ID no 19)
25	CTC TCC CTA CCG GCT (selected from positions 1734-1740 in Figure 1H),	(Seq ID no 20)
	GAT ATT CCG GTC CCC (selected from positions 1967-1976 in Figure 1H),	(Seq ID no 21)
	ACT CCG CCC CAA CTG (selected from positions 2403-2420 in Figure 1H),	(Seq ID no 22)
	CTG TCC CTA AAC CCG (selected from positions 2457-2488 in Figure 1I),	(Seq ID no 23)
	TTC GAG GTT AGA TGC (selected from positions 2457-2488 in Figure 1I),	(Seq ID no 24)
30	GTC CCT AAA CCC GAT (selected from positions 2457-2488 in Figure 1I),	(Seq ID no 25)
	GGT GCA CCA GAG GTT (selected from positions 2952-2956 in Figure 1I),	(Seq ID no 26)
	CTG GCG GGA CAA CTG (selected from positions 2966-2969 in Figure 1J),	(Seq ID no 27)
	TTA TCC TGA CCG AAC (selected from positions 3000-3003 in Figure 1J),	(Seq ID no 28)
	GAC CTA TTG AAC CCG (selected from positions 3097-3106 in Figure 1J),	(Seq ID no 29)
35		
	GAA GAG ACC TTT CCG (selected from positions 76-79 in Figure 2A),	(Seq ID no 30)
	CAC TCG AGT ATC TCC (selected from positions 98-101 in Figure 2A),	(Seq ID no 31)
	ATC ACC CAC GTG TTA (selected from positions 136-136 in Figure 2A),	(Seq ID no 32)
40	GCA TCC CGT GGT CCT (selected from positions 194-201 in Figure 2B),	(Seq ID no 33)
40	CAC AAG ACA TGC ATC (selected from positions 194-201 in Figure 2B),	(Seq ID no 34)
	TAA AGC GCT TTC CAC (selected from positions 222-229 in Figure 2B).	(Seq ID no 35)
	GCT CAT CCC ACA CCG (selected from position 242 in Figure 2B),	(Seq ID no 36)

AAA CAG TCG CTA CCC (selected from positions 2025-2027 in Figure 4I),

CCT TAC GGG TTA ACG (selected from positions 2131-2132 in Figure 4J).

GAG ACA GTT GGG AAG (selected from positions 2252-2255 in Figure 4J),

TGG CGT CTG TGC TTC (selected from positions 2396-2405 in

(Seq ID no 72)

(Seq ID no 73)

(Seq ID no 74)

20

30

35

	Figure 4J and Figure 4K),	(Seq ID no 75)
	CGA CTC CAC ACA AAC (selected from positions 2416-2420 in Figure 4K),	(Seq ID no 76)
	GAT AAG GGT TCG ACG (selected from positions 2474-2478 in Figure 4K),	(Seq ID no 77)
	ATC CGT TGA GTG ACA (selected from position 2687 in Figure 4K),	(Seq ID no 78)
5	CAG CCC GTT ATC CCC (selected from position 2719 in Figure 4K),	(Seq ID no 79)
	AAC CTT TGG GAC CTG (selected from position 2809 in Figure 4L),	(Seq ID no 80)
	TAA AAG GGT GAG AAA (selected from positions 3062-3068 in Figure 4L),	(Seq ID no 81)
	GTC TGG CCT ATC AAT (selected from positions 3097-3106 in Figure 4L),	(Seq ID no 82)
10	AGA TTG CCC ACG TGT (selected from positions 135-136 in Figure 5A),	(Seq ID no 83)
	AAT CCG AGA AAA CCC (selected from positions 472-475 in Figure 5A),	(Seq ID no 84)
	GCA TTA CCC GCT GGC (selected from positions 1136-1144 in Figure 5B),	(Seq ID no 85)
	TTA AAA GGA TTC GCT (selected from positions 1287-1292 in Figure 5B),	(Seq ID no 86)
	AGA CCC CAA TCC GAA (selected from position 1313 in Figure 5B),	(Seq ID no 87)
15	GAC TCC GAC TTC ATG (selected from position 1334 in Figure 5B),	(Seq ID no 88)
	GTC TTT TCG TCC TGC (selected from positions 2568-2569 in Figure 6),	(Seq ID no 89)
	GTC TTA TCG TCC TGC (selected from positions 2568 in Figure 6),	(Seq ID no 90)
	GTC TTC TCG TCC TGC (selected from positions 2568 in Figure 6),	(Seq ID no 91)
- 20	GTC TTG TCG TCC TGC (selected from positions 2568 in Figure 6),	(Seq ID no 92)
	GTC TAT TCG TCC TGC (selected from positions 2568 in Figure 6),	(Seq ID no 93)
	GTC TCT TCG TCC TGC (selected from positions 2568 in Figure 6),	(Seq ID no 94)
	GTC TGT TCG TCC TGC (selected from positions 2568 in Figure 6),	(Seq ID no 95)
25	TTG GCC GGT GCT TCT (selected from positions 452 in Figure 7),	(Seq ID no 96)
	TTG GCC GGT ACT TCT (selected from positions 452 in Figure 7),	(Seq ID no 97)
	TTG GCC GGT CCT TCT (selected from positions 452 in Figure 7),	(Seq ID no 98)
	TTG GCC GGT TCT TCT (selected from positions 452 in Figure 7),	(Seq ID no 99)
	ACC GCG GCT GCT GGC (selected from positions 473-477 in Figure 7),	(Seq ID no 100)
30	ACC GCG GCT ACT GGC (selected from positions 473 in Figure 7),	(Seq ID no 101)
	ACC GCG GCT CCT GGC (selected from positions 473 in Figure 7), or	(Seq ID no 102)
	ACC GCG GCT TCT GGC (selected from positions 473 in Figure 7),	(Seq ID no 103)
	CGG CAG CTG GCA CGT (selected from positions 474 in Figure 7),	(Seq ID no 104)
	CGG CCG CTG GCA CGT (selected from positions 474 in Figure 7),	(Seq ID no 105)
35	CGG CTG CTG GCA CGT (selected from positions 474 in Figure 7),	(Seq ID no 106)
	CGT ATT ACC GCA GCT (selected from positions 477 in Figure 7),	(Seq ID no 107)
	CGT ATT ACC GCC GCT (selected from positions 477 in Figure 7),	(Seq ID no 107)
	CGT ATT ACC GCT GCT (selected from positions 477 in Figure 7),	(Seq ID no 109)
40	TTC CTT TGA GTT TTA (selected from positions 865-866 in Figure 7),	(Seq ID no 110)
40	TTC CTT TAA GTT TTA (selected from positions 865 in Figure 7),	(Seq ID no 111)
	TTC CTT TCA GTT TTA (selected from positions 865 in Figure 7),	(Seq ID no 112)
	TTC CTT TTA GTT TTA (selected from positions 865 in Figure 7),	(Seq ID no 113)
	TTC CTT AGA GTT TTA (selected from positions 866 in Figure 7),	(Seq ID no 114)

	WO 98/15648	68	PCT/DK97/00425
	TTC CTT CGA GTT TTA (selected from position	ons 866 in Figure 7),	(Seq ID no 115)
	TTC CTT GGA GTT TTA (selected from position	ons 866 in Figure 7),	(Seq ID no 116)
	CAT GTG TCC TGT GGT		(Seq ID no 117)
	CGT CAG CCC GAG AAA		(Seq ID no 118)
,	CAC TAC ACA CGC TCG		(Seq ID no 119)
	TGG CGT TGA GGT TTC and		(Seq ID no 120)
	AAC ACT CCC TTT GGA		(Seq ID no 123)
	and a mixture of such probes.		
,	23. Peptide nucleic acid probes according	to claim 22, wherein the Qs	of adjacent moleties
	are selected so as to form the following su		•
	TCA CCA CCC TCC TCC		(Seq ID no 6)
;	CCA CCC TCC TCC		(modified Seq ID no 6)
	ACT ATT CAC ACG CGC		(Seq ID no 8)
	CCA CAC CCA CCA CAA		(Seq ID no 12)
	AAC TOO ACA COO COG		(0 ID 40)

	104 004 000 100 100	(Sed ID to 6)
15	CCA CCC TCC TCC	(modified Seq ID no 6)
٠.	ACT ATT CAC ACG CGC	(Seq ID no 8)
	CCA CAC CCA CCA CAA	(Seq ID no 12)
	AAC TCC ACA CCC CCG	(Seq ID no 16)
	ACT CCA CAC CCC CGA	(Seq ID no 17)
20	ACT CCG CCC CAA CTG	(Seq ID no 22)
	CTG TCC CTA AAC CCG	(Seq ID no 23)
	TTC GAG GTT AGA TGC	(Seq ID no 24)
	GTC CCT AAA CCC GAT	(Seq ID no 25)
	GAC CTA TTG AAC CCG	(Seq ID no 29)
25		
	GCA TCC CGT GGT CCT	(Seq ID no 33)
	CAC AAG ACA TGC ATC	(Seq ID no 34)
	GGC TTT TAA GGA TTC	(Seq ID no 40)
	24724	:
30	GAT CAA TGC TCG GTT	(Seq ID no 44)
	CGA CTC CAC ACA AAC	(Seq ID no 76)
	GCA TTA CCC GCT GGC	(Seq ID no 85)
35	GTC TTA TCG TCC TGC	(Seq ID no 90)
	GTC TTC TCG TCC TGC	(Seq iD no 91)
	GTC TTG TCG TCC TGC	(Seq ID no 92)
	GTC TAT TCG TCC TGC	(Seq ID no 93)
	GTC TCT TCG TCC TGC	(Seq ID no 94)
40	GTC TGT TCG TCC TGC	(Seq ID no 95)
	AAC ACT CCC TTT GGA	(Seq ID no 123)

(OK 747/modified Seq ID no 93)

	CAT GTG TCC TGT GGT			
	CGT CAG CCC GAG AAA	(Seq ID no 117)		
	OUT ONG COL GAG AAA	(Seq ID no 118)		
5	CAC TAC ACA CGC TCG,	(Seq ID no 119)		
	TGG CGT TGA GGT TTC	(Seq ID no 120)		
	and a minture of such such			
	and a mixture of such probes.			
10	24. Peptide nucleic acid probes according to claim 22 or 23 selected from			
	Lys(Flu)-Lys(Flu)-TCA CCA CCC TCC TCC-NH2	(OK 446/modified Seq ID no 6)		
	Lys(Flu)-Lys(Flu)-CCA CCC TCC TCC-NH ₂	(OK 575/modified Seq ID no 6)		
	Lys(Flu)-Lys(Flu)-ACT ATT CAC ACG CGC-NH₂	(OK 447/modified Seq ID no 8)		
15	Lys(Fiu)-ACT ATT CAC ACG CGC-NH ₂	(OK 688/modified Seq ID no 8)		
	Lys(Flu)-Lys(Flu)-CCA CAC CCA CCA CAA-NH2	(OK 448/modified Seq ID no 12)		
	Lys(Flu)-Lys(Flu)-AAC TCC ACA CCC CCG-NH ₂	(OK 449/modified Seq ID no 16)		
	Lys(Flu)-Lys(Flu)-ACT CCA CAC CCC CGA-NH ₂	(OK 309/modified Seq ID no 17)		
•	Lys(Flu)-Lys(Flu)-ACT CCG CCC CAA CTG-NH2	(OK 450/modified Seq ID no 22)		
20	Lys(Flu)-Lys(Flu)-CTG TCC CTA AAC CCG-NH₂	(OK 305/modified Seq ID no 23)		
	Lys(Flu)-Lys(Flu)-TTC GAG GTT AGA TGC-NH₂	(OK 306/modified Seq ID no 24)		
	Lys(Flu)-TTC GAG GTT AGA TGC-NH₂	(OK 682/modified Seq ID no 24)		
	Lys(Flu)-Lys(Flu)-GTC CCT AAA CCC GAT-NH₂	(OK 307/modified Seq ID no 25)		
	Lys(Flu)-GTC CCT AAA CCC GAT-NH ₂	(OK 654/modified Seq ID no 25)		
25	Lys(Flu)-GAC CTA TTG AAC CCG-NH₂	(OK 660/modified Seq ID no 29)		
	Lys(Flu)-Lys(Flu)-Gly-GCA TCC CGT GGT CCT-NH₂	(OK 223/modified Seq ID no 33)		
	Lys(Flu)-Lys(Flu)-CAC AAG ACA TGC ATC-NH₂	(OK 310/modified Seq ID no 34)		
	Lys(Flu)-CAC AAG ACA TGC ATC-NH₂	(OK 655/modified Seq ID no 34)		
30	Lys(Flu)-GGC TTT TAA GGA TTC-NH2	(OK 689/modified Seq ID no 40)		
	Lys(Rho)-GGC TTT TAA GGA TTC-NH₂	(OK 702/modified Seq ID no 40)		
	Flu-β-Ala-β-Ala-GAT CAA TGC TCG GTT-NH ₂	(OK 624/modified Seq ID no 44)		
	Flu-β-Ala-β-Ala-CGA CTC CAC ACA AAC-NH ₂	(OK 612/modified Seq ID no 76)		
35	•	(**************************************		
	Flu-β-Ala-β-Ala-GCA TTA CCC GCT GGC-NH₂	(OK 623/modified Seq ID no 85)		
	Lys(Flu)-GTC TTT TCG TCC TGC-NH₂	(OK 745/modified Seq ID no 89)		
	Lys(Rho)-GTC TTA TCG TCC TGC-NH ₂	(OK 746/modified Seq ID no 90)		
40	Lys(Rho)-GTC TTC TCG TCC TGC-NH ₂	(OK 746/modified Seq ID no 91)		
	Lys(Rho)-GTC TTG TCG TCC TGC-NH₂	(OK 746/modified Seq ID no 92)		

Lys(Rho)-GTC TAT TCG TCC TGC-NH₂

WO 98/15648

PCT/DK97/00425

Lys(Rho)-GTC TCT TCG TCC TGC-NH2

(OK 747/modified Seq ID no 94) (OK 747/modified Seq ID no 95)

Lys(Flu)-AAC ACT CCC TTT GGA-NH,

Lys(Rho)-GTC TGT TCG TCC TGC-NH2

(OK 749/modified Seq ID no 123)

5

wherein Flu denotes a 5-(and 6)-carboxyfluoroescein label and Rho denotes a rhodamine label,

70

and a mixture of such probes.

10 25. Use of a peptide nucleic acid probe according to any one of claims 1 to 24 or a mixture thereof for detecting a target sequence of one or more mycobacteria optionally present in a sample.

- 26. Use of a peptide nucleic acid probe or a mixture thereof according to claim 25 for detecting a target sequence of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC), in particular a target sequence of M. tuberculosis.
- 27. Use of a peptide nucleic acid probe or a mixture thereof according to claims 25 for detecting a target sequence of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex, in particular a target sequence of one or more mycobacteria of the Mycobacterium avium Complex.
- 28. Method for detecting a target sequence of one or more mycobacteria optionally present in a sample comprising

25

15

20

(1) contacting any rRNA or rDNA present in said sample with one or more peptide nucleic acid probes according to any one of claims 1 to 24 or a mixture thereof under conditions, whereby hybridisation takes place between said probe(s) and said rRNA or rDNA, and

- (2) observing or measuring any formed detectable hybrids, and relating said observation or measurement to the presence of a target sequence of one or more mycobacteria in said sample.
- 29. Method according to claim 28 for detecting a target sequence of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC), in particular a target sequence of M. tuberculosis.
 - 30. Method according to claim 28 for detecting a target sequence of one or more

5

mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex.

- 31. Method according to any one of claims 28 to 30, wherein the hybridisation takes place in situ.
- 32. Method according to any of of claims 28 to 30, wherein the hybridisation takes place in vitro.
- 33. A method according to any one of claims 28 to 32,
- 10 characterised in that a signal amplifying system is used for measuring the resulting hybridisation.
 - 34. Method according to any one of claims 28 to 33, wherein the sample is a sputum sample.
- 35. Kit for detecting a target sequence of one or more mycobacteria, in particular a target sequence of one or more mycobacteria of the Mycobacterium tuberculosis Complex (MTC), in particular a target sequence of M. tuberculosis, and/or for detecting a target sequence of one or more mycobacteria other than mycobacteria of the Mycobacterium tuberculosis Complex (MOTT), in particular a target sequence of one or more mycobacteria of the Mycobacterium avium Complex,
 - c h a r a c t e r i s e d in that said kit comprises at least one peptide nucleic acid probe according to any one of claims 1 to 24, and optionally a detection system with at least one detecting reagent.
- 36. Kit according to claim 35,c h a r a c t e r i s e d in that it further comprises a solid phase capture system.

		130	140	150	16	=
1093	GGGGAAAC	CCAGCACGA	GTGATGTCGT	CTACCCGCA	TET	M.tuberculosis
422	GGGGGAAG	CCAGCACGA	GTGATGTCGT	TRACCCGRA	TCT	M.avium
422	GGGGGAAC	CCCAGCACGA	GTGATGTCGT	ShTACCCGHA	TCT	M.paratuberc
507	GGGGGAAC	CCCGGCACGA	GTGATGTCGT(STOACCOAAC	GCT.	M.phlei
432	GGGGAAAC	CCAACACGA	STCACCTCCT	SHTACCCGHA	TCT	M.lenrae
207	GGGGAAAC	CCAGCACGA	GTAATGTCGT	Shitacccghia	TCT	M.gastri
150	GGGGAAAC	CCAGCACGÁ	GTGATGTCGT	STRACCCGCA	TCT	M.kansasii
2588	GGGGAAAC	CCCGGCACGA	STGATGTCGT	STOACCAGEC	д ст	M.smegmatis
		_			_	· · · · · · · · · · · · · · · · · · ·
				_		
	100 V					
		210	220	230	240	
1172	CATCTCAG	TACCCGTAG	FAGGAGAAAAC	CAATTGTGAT	rcc	M.tuberculosis
501	CATCTCAG	TACCCGTAG	GAGAAGAAAAC	CAATTGTGAT'	TCC	M.avium
501	CATCTCAG	TACCCGTAG	BAGAAGAAAAC	CAATTGTGAT'	rcc	M.paratuberc.
586	CATCTCAG	TACCCGTAG	AGAAGAAAAC	CAATTGTGAT'	rcc	M.phlei
511	CATCTCAG	TACCCGTAG	AGAAGAAAA	AATTGTGAT'	rcc	M.leprae
286	CATCTCAG	TACCCGTAG	FAGAAGAAAAC	AAAAGTGAT!	rcc	M.gastri
229	CATCTCAG	TACCCGTAG	SAGNAGAAAAC	AAAAGTGAT!	rcc	M.kansasii
2667	CATCTCAG	THCCCGTAG	SAAGAAAAC	AAMTGTGAT!	rcc	M.smegmatis
		•		•		
		· · · · · · · · · · · · · · · · · · ·				
		330	340	350	360)
1289	TGTGGGAG	GATATGTCT	CAGCGCTACC	CGGCTGAGA-	-GG	M.tuberculosis
617	TGTGGGAT	TGATATGTCT	CAGCICTACC	TGGCTGAGG	-GG	M.avium
617	TGTGGGAT	TGATATGTC1	CAGCICTACC	TIGGCTGAGG	-GG	M.paratuberc.
703	TGTGGGGC	CTGTGTGTC	CAMCGRCCGC	CGGCGAAGG	בותי	M phloi
629	TGTGGGAT	TGGTATGTCT	CARCICTACO	TIGGITGAGG	-GG	M.leprae
404	TGTGGGAT	'GĀTAĢGTC'I	CAGCICTACC	CGGCTGAGG	-GG	M.gastri
347	TGTGGGAT	'dgatadgtc'	CAGCICTACC	CGGCTGAGG	-GG	M.leprae M.gastri M.kansasii
2785	TGTGGGAC	CTATOTUTO	CECCTCTACC	Десте⊟с¤с	GG	M.smegmatis
						-

Figure 1A

		370	380	390	40	n
1327	CAGTCAG	AAAGTGTCGT	GGTTAGCGGA	AGUGGCCUGG		_
656	HAGTCAG	AAAGTGTCGT	GGTTAGCGGA	AGIGGCCIGG	CAT	M.tuberculosis M.avium
656	TAGTCAG	AAAGTGTCGT	GGTTAGCGGA	AGTGGCCTGG AGTGGCCTGG	H _Z	M. paratuberc.
742	TAGTGAT	AAAGCAGTGT	GGTTAGGTGA	AGTGGCCTGG	GAT	M.phlei
668	TAGTCAG	AAAGTGCCGT	GGTTAGCGGA	AATGGCCTGG	GAT	M.leprae
443	CAGTCAG	aaagtgtcgt	GGTTAACGGA	AGTGGCCTGG	GAT	M.gastri
386	CAGTCAG	AAAGTGTCGT	GGTTANCGGA	В СТСССТСС	ርውጥ	M kangagii
2823	CAGTGAG.	aaaAtgtiigt	ggttagcgga	AATGGCTTGG	GAT	M.smegmatis
						J
		450	460	470	480	
1406	CGGCACC	TGCCTAGTAT	CAATTCCCGA	GTAGCAGCGG	GCC	M.tuberculosis
735	CGGCACC'	TGCCTTATAT	Caacacccga	GTAGCAGCGG	GCC	M.avium
735	CGGCACC'	TGCCTTATAT	Caacacccga	GTAGCAGCGG	GCC	M. paratuberc
820	TECTECO	-GCTGTCACA	GG-⊣TCCCGA	GTAGCAGCGG	GCC	M.phlei
747	IGGCACC'	igcct igtat	CAATTCCCGA	GTAGCAGCGG	GCC	M.leprae
522	CGGCACC'	PGCCTIGTAT	CAATTCCCGA	GTAGCAGCGG	GCC	M.gastri
465	CGGCACC	recettietat	CAATTCCCGA	GTAGCAGCGG	GCC	M.kansasii
2902	CGACGIC	relictigate	GTGTTCCCGA	GTAGCAGCGG	GCC	M.smegmatis
						
		490	500	510	520	
1446	CGTGGAA	TCCCTGTGA	ATCC CCGGG	ACCACCGGT	AAG	M.tuberculosis
775	CGTGGAA	PCTGTGA:	atchecces:	ACCACCCGGT:	AAG	M.avium
775	CGTGGAA	rchectetea:	atchecege:	ACCACCCGGT:	AAG	M.paratuberc.
857	CGTGGAA	PCIGCTGTGA	atchecces:	ACCACCCGGT:	AAG	M.phlei
787	CGTGGAA	rchgctgtga:	ATCHGCCGGG:	ACCACCCGGT:	AAG	M.leprae
562	CGTGGAA	PCIECTGTGA	ATC GCCGGG	ACCACCGGT	AAG	M.gastri
505	CGTGGAAT	rchectetem	ATCIIGCCGGG:	ACCACCCGGT	AAG	M.kansasii
2942	CGTGGAA	rchectetem	ATC <u>H</u> GCCGGG	ACCACCCGGT	AAG	M.smegmatis

Figure 1B

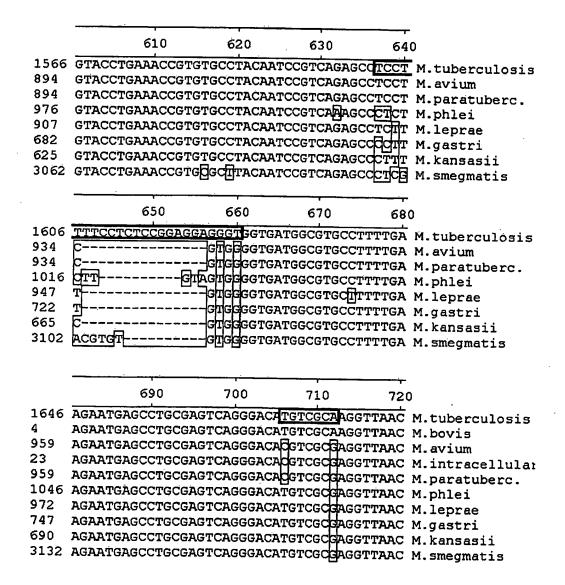
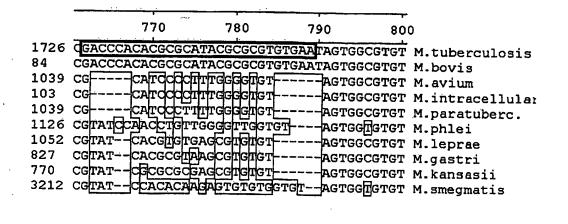


Figure 1C



	970	980	990	1000	
1926	ATTTAGGTGCAGC	TTGCGTGGTTC	COGCGGAG	GTAGAG M.tuberculo	sis
1228	ATTTAGGTGCAGC	TTGCGTGGTTC;	CCACGGAG	GTAGAG M.avium	
1228	ATTTAGGTGCAGC	TTGCGTGGTTC	ACCACGGAG	GTAGAG M.paratuber	c.
1322	ATTTAGGTGCAGC	STOGCATGITTC	TATCGGAG	GTAGAG M.phlei	
1244	ATTTAGGTGCAGC	TTGCGTGGTTC	ACCAICGGAG	GTAGAG M.lenrae	
1019	ATTTAGGTGCAGC	TTGCGTGTTTC1	CCACGGAG	GTAGAG M.gastri	
962	ATTTAGGTGCAGC	TTGCGTGHTTC	CCACGGAG	GTAGAG M.kansasii	
3408	ATTTAGGTGCAGC	ridechreihre	TECCGGAG	GTAGAG M. smegmatis	

	105	50 1	060	1070	1080
2005	CAGCCAAACT	CCGAATGC	G-TGGTG-	TA-AAGCGTG	GCA M.tuberculosis
1307	CAGCCAAACT	CCGAATGC	CG-TGGTG-	TAAAAGCGTG	GCA M.avium
1307	CAGCCAAACT	CCGAATGC	CG-TGGTG-	TANAAGCGTG	GCA M.paratuberc
1401	CAGCCAAACT	CCGAATGC	GATAAG'	TGALAGTIGTG	GCA M.phlei
1323	CAGCCAAACT	CCGAATGC	G-TGGTI-	TAMAAGCGTG	GCA M.leprae
1098	CAGCCAAACT	CCGAATGC	CG-TGGTG-	TATA-GCGTG	GCA M.gastri
1041	CAGCCAAACT	CCGAATGC	CG_TGGTG-	TATA GCGTG	GCA M.kansasii
3486	CAGCCAAACT	CCGAATGC	CGGTAAGGC	CALAGRAGIIG CG	GAA M. smeamatis

Figure 1D

	1130	1140	1150	1160
082	ACAGCCCAGATCGC	CGCTAAGGCC	CCCAAGCGT	TGCTA M.tuberculo
385	ACAGCCCAGATCGCC	CGGCTAAGGCC	CONARGCGTO	TGCTA M.avium
385	ACAGCCCAGATCGC	CGGCTAAGGCC	CCTAAGCGTG	TGCTA M.paratuber
479	ACAGCCCAGATCGCC	CGGCTAAGGCC	CCTAAGCGTG	TGCTA M.phlei
401	ACAGCCCAGATCGC	CGGCTAAGGCC	CCTAAGCGTG	TGCTA M.leprae
175	ACAGCCCAGATCGCC	CGGCTAAGGCC	CCAAAGCGTG	TGCTA M.gastri
118	ACAGCCCAGATCGCC	GGCTAAGGCC	CCALAGCGTC	TGCTA M.kansasii
566	ACAGCCCAGATCGC	GGTTAAGGCC	CCHAAGCGT	TGTTA M. smegmatis

	1290	1300	1310	1320
241 CTCAP	AGCACACCGC	CGAAGCCGCGG	CACATCCACC	TTGT- M.tuberculos
544 CTCAP	AGCACACCGC	CGAAGCCGCGG	CACATICATO	TT TA M. avium
544 CTCAP	AGCACACCGC	CGAAGCCGCGG	CACATICATO	TT TW M. paratubero
538 CTCAP	AGCACACCGC	CGAAGCCGCGG	CAF-ATCAGO	
660 CTCAP	AGCACACCGC	CGAAGCCGCGG	CACATICACO	
334 CTCAP	AGCACACCGC	CGAAGCCGCG	CAFACC	
277 CTCAP	AGCACACCGC	CGAAGCCGCG	CAACC	GCA M.kansasii
726 ПТСАЛ	AGCACACCGC	CGAAGCCGCGG	AAGCCAAC	GTHTG M. smegmatis
_				
	· · · · · · · · · · · · · · · · · · ·	· · · · · · · · · · · · · · · · · · ·		
<u> </u>	1330	1340	1350	1360
:80 -GGTG				•
80 <u>- GGTG</u>	GGTGTGGGT	AGGGGAGCGTC	CCTCATTCAG	GCGAAG M.tuberculos
83 CGGTG	GETGEGGTA	AGGGGAGCGTC	CCTCATTCAG	GCGAAG M.tuberculos GCGAAG M.avium
83 CGGTG	GETGEGETA FETGEGETA FETGEGETA	AGGGGAGCGTC AGGGGAGCGTC	CCTCATTCAG CCCCATTCAG	CGAAG M.tuberculos CGAAG M.avium CGAAG M.paratuberc
683 CGGTG 683 CGGTG 576 TGGCT	egteregti eptetegeti eptetegeti egtetegeti	AGGGGAGCGTC AGGGGAGCGTC AGGGGAGCGTC AGGGGAGCGTC	CCTCATTCAG CCCCATTCAG CCCCATTCAG	CGAAG M.tuberculos CGAAG M.avium CGAAG M.paratuberc
683 CGGTG 683 CGGTG 576 TGGCT 500 GGGTG	egtgreet epteteget egteteget geteteget epteteget	AGGGAGCGTC AGGGGAGCGTC AGGGGAGCGTC AGGGGAGCGTC AGGGGAGCGT[CCTCATTCAG CCCCATTCAG CCCCATTCAG CTGCATGCGG CCTCATTCAG	CGAAG M.tuberculos CGAAG M.avium CGAAG M.paratuberc TGAAG M.phlei CGAAG M.leprae
683 CGGTG 683 CGGTG 576 TGGCT 500 GGGTG	egtgreetr egtetegetr egtetegetr getetegetr egtetegetr	AGGGGAGCGTC AGGGGAGCGTC AGGGGAGCGTC AGGGGAGCGTC	CCTCATTCAG CCCCATTCAG CCCCATTCAG CTGCATCCGG CCTCATTCAG	CGAAG M.tuberculos CGAAG M.avium CGAAG M.paratuberc CGAAG M.phlei CGAAG M.leprae CGAAG M.gastri

Figure 1E

		1				
	13		1380	1390	140	
2319	CCACCGGGTG	SACCGGTG	GTGGAGGGT	GGGGGAGTGA	CAAT	M.tuberculosis
1623	CT-CCGGGTG	ZACCGĠTG	CTCCNCCCT	CCCCCNCMCX	CYVUI	M. cuberculosis
1623	CT-CCGGTG	znfficecme	GMCCPCCCM	CCCCCRCRC	GAAT	M.avium
	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	271100010	GIGGAGGGI CECCECCE	GGGGGAGTGA	GAAT	M.paratuberc.
1640	соссейств	2 ATICGGIG	GTGGAGGGT STSSSSSSSSS	GIIGGGAGTGA	GAAT	M.phlei
		ACCGGTG	GTGGAGGGT	ggggaatga	GAAT	M.leprae
1402		SACCGGTG	GTGGAGGAT	gggggagtga	GAAT	M.gastri
1345	cliecceeere	ACCGGTG	GTGGAGGAT	<u>G</u> GGGGAGTGA	GAAT.	M.kansasii
3796	COGCCGAGTE	ATCGAGTG	GTGGAGGGT	GIGGGAGTGA	GAAT	M.smegmatis
						-
	2.0	10	1 100			
	14		1420	1430	144	•
2359	GCAGGCATGA	GTAGCGA	CAAGGCAAG	TGAGAACCTT	GCCC	M.tuberculosis
1662	GCAGGCATGA	agtagcga	ĪAAGGCAAG'	TGAGAACCTT	GCCC	M. avium
1662	GCAGGCATGA	GTAGCGA	IAAGGCAAG'	TGAGAACCTT	GCCC	M.paratuberc.
1756	GCAGGCATGA	GTAGCGA	TAAGGCAAG	TGAGAACCTT	Mccc	M phlai
1680	GCAGGCATGA	GTAGCGA	TAAGGCAAG'	TGAGAACCTT	ECCC	M lenne
1442	GCAGGCATGA	GTAGCGA	DAAGGCAAG	™GDGDDCC™™	GCCC	M. reprae
1385	GCAGGCATGA	GTAGCGA	DARGCAAG'	™CDCDACOTI TCDCDACCOTU	GCCC	M.kansasii
3836	GCAGGCATGA	CTACCCA	rincecane:	TCD CDD CCDM	3222	M.smegmatis
5050	CONDUCTION	ro raccan	I INGGCAAG.	I GAGAACCTT	gece	M.smegmatis
				-		
		····				
	15	70 .	1580	1590	160	0
2519	CCCCGTGAC	GAATCA-	CGGTACTA	ACCACCCAAA	ACCG	M.tuberculosis
1821	CGICCOTGAI	GAATCA-	CCCTACTA			M origina
1821	Cencoorean	GDATCA-	こしてには山かしむかっ	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	200	M.paratuberc.
	CGICCOTGAT	Charche		っししれしししAAA	2005	m.paratuberc.
1840	CCCCCCCCC	CAMBON C	TI TO IBCTA	ACCACCCAAA	4001	m.pnrei
1602	CGCCCGTGAT	GAATCA-	CCCCTACTO	ACCACCCAAA	ACCG	M.leprae
1545	CGCCCGTGAT	GAATCA-	CGGTACTA	ACCACCCAAA	ACCG	M.gastri
	CGCCCGTGAT	GAATCA-(SCGGTACTAI	ACCACCCAAAI	ACCG	M.kansasii
3996	Cellcourgui	gaatca-(GGGTACTAI	ACCAIICCAAA	ACCE!	M.smegmatis

Figure 1F

					
	161				~
2558	GAT-CGATCA	C-TCCCCTTCG	GGGG-TGTGG	AGTTC-TGG	M.tuberculosis
1860	GAT-CGACCA	T-TCCCCTTCG	GGGGC-GTGG	GATIT-DGG	M. avium
1860	GAT-CGACCA	T-TCCCCTTCG	ggggc-gtgg	GATTI-CGG	M. paratuhero
1955	GGG-CGATC-	−AITCCITTCG	GGGH−HGTGĀ	GGTTG-GG	Monhlei
1879	GAT-CGACCA	TATCCCCTTCG	GGGGTATGG	AGGTTF-OGG	M.leprae
1641	GAT-CGATCA	C-TCCCCTTCG	GGGGA-GTGG	AGGTC-TGG	M.gastri
1584	GAT-CGATCA	C-TCCCCTTCG	GGGC-GTGG	AGGTC-TGG	M. kansasii
4035	ACCGTGACCG	CACCITICG	GGGF-TGTGG	GTTGGTGG	M.smegmatis
			ب ب		***************************************
	1.65	0 1666	1.070		_
	165				
2594	GGCTGCGTGG	SAACTTCGCTG	GTAGTAGTCA	AGCGAAGGG	M.tuberculosis
1896	GGCTGCGTGG	SAICTTCGCTG:	GTAGTAGTCA	AGCAARIGGG	M. avium
1896	GGCTGCGTGG	SACCTTCGCTG	GTAGTAGTCA	AGCHANGGG	M.paratuberc.
1986	GGCTGCGTGG	SAICCG-GIGG	GTAGTAGTCA	AGCGATGGG	M.phlei
1917	GGCTGCGTGG	SAACTTCGTTG	GTAGTAGTCA <i>I</i>	AGCGATGGG	M.leprae
1677	GGCTGCGTGG	AGCTTCGCTG	GTAGTAGTCA	GCGATGGG	M.gastri
1620	GGCTGCGTGG	agqcttcgctg:	STAGTAGTCA <i>I</i>	AGCGAITGGG	M.kansasii
4071	GGCTGCATGG	FACTTCGTTG	STAGTAGTCA	GCGATGGG	M.smegmatis
					
	169			1720	
2634	-GTGACGCAGG	BAAGGTAGCCG'	TACCAGTCAGT	GGTAACA-	M.tuberculosis
1936	-GTGACGCAGG	Baagg agccg'	PACCAGTCAG	GGTAAIIA-	M.avium
1936	-GTGACGCAGG	Faaggdagccg'	PACCAGTCAGT	GGTAATA-	M.paratuberc.
2025	-GTGACGCAGG	Baaggtagccg'	TACCAGTCAGT	GGTAATA-	M.phlei
1957	-GTGACGCAGG	SAAGGTAGCCG'	FACCAGTCAG1	GGTAAHA-	M.leprae
1717	-GTGACGCAGG	Baaggdagccg'	TACCAGTCAGT	GGTAAHA-	M.gastri
1660	-GTGACGCAGG	SAAGGBAGCCG!	FACCAGTCAGT	GGTAATA-	M.kansasii
4111	-GTGACGCAGG	BAAGGTAGCCG'	facc e gtcagi	GGTAATA-	M.smegmatis

Figure 1G

1730	1740	1750	1760
2672 -CTGGGGCAAGCCGG 1974 -CTGGGGCAAGCCGG 1974 -CTGGGGCAAGCCGG 2063 -CGGGGGTAAACCGG 1995 -CTGGAGCAAGCCAGG 1755 -CTGGGGCAAGCCAGG 1698 -CTGGGGCAAGCCAGG 4149 -CGGGGGTAAGCCTGG	Pagada Pagada Pagada Pagada Pagada Pagada Pagada Pagada Pagada	CGATAGGCAAA CGATAGGCAAA CGATAGGCAAA CGATAGGCAAA CGATAGGCAAA CGATAGGCAAA	TCCGT M.avium TCCGT M.paratuberc. TCCGT M.phlei TCCGT M.leprae TCCGT M.gastri TCCGT M.kansasii

	1970	1980	1990	2000)
2908	AGGGGGACCGGAATA'	ICGTGAACA	CCCTTGCGGTG	GGAGC	M.tuberculosis
2208	AGGGGGCCGGAATA	CGTGAACA	CCCTTGCGGTG	GGAGC	M.avium
2208	AGGGGGCCGGAATA	CGTGAACA	CCCTTGCGGTG	GGAGC	M. paratuberc
2298	AGGGGGACCCACGTA	CGTGAGGG	CTCTTGCGGGG	sDage -	M.phlei
2231	AGGGGGGCCGGAATA!	ICGTGAACA	CCCTTGCGGTG	GAGC	M.leprae
1910		_			M.gastri
1934	AGGGGGACCGGAATA	CGTGAACA	CCCTTGCGGTG	GAGC	M.kansasii
4385	AGGGGGACCCACATE	CGTGTAAG	CCITTIFICGGCC	CAAGC	M.smegmatis

		•					· ·
		2410	. 24		2430	244	· •
3345	ACCTCGA	CGCCAG'	TGGGGC	GAGTC	TTGTTGAA	ATACC	M.tuberculosis
284	ACCTCGA	CGCCAG'	TGGGGC	GGAGTC	TTGTTGAA	ATACC.	M.bovis
2645	GCACAGA	CGCCAG'	rriigiigi	GGAGTC	TTGTTGAA	ATACC	M. avium
393	ATACAGA	CGCCAG:	rthghāi	GGAGTC	TTGTTGAA	ATACC	M.intracellulare
2645	GCACAGA	.CGCCAG	TIGIGI	GGAGTC	TTGTTGAA	ATACC	M.paratuberc.
2737	GCTCGGA	CGCCAG	TOGGGI	GGAGTC	TTGTTGAA	ATACC	M.phlei
2668	ACTITCGA	CGC[[AG:	rTGGGG	GGAGTC	TTGTTGAA	ATACC	M.leprae
1910							M.gastri
2372	ACCTCAA	CGCCAG'	rrggggfi	GGAGTC	TTGTTGAA	ATACC	M. Kansasii
4822	GCTCACA	CGCCAG	GIGGGI	GGAGTC	TTGTTGAA	ATACC	M.smegmatis

Figure 1H

						
		2450	2460	2470	2480	
3385	ACTCTGA	TCGTATTG	GCATCTAAC	CTCGAACCCT	GAATC M.tuberculosis	=
324	ACTCTGA	TCGTATTG	GCATCTAAC	CTCGAACCCT	GAATC M boyis	•
2685	ACTCTGA	TCGTATTG	ACADCTAAC	GTCGAACCCT	TATC M.avium	
433	ACTCTGA	TCGTATTG	PICACCTAAC	STCGAACCCT	TATC M.intracellular	are
2685	ACTCTGA	TCGTATTG	SACACCTAAC	GTCGAACCCT	HTATC M paratubara	
2777	ACTCTGA	TCGTATTG	GCCTCTAAC	CTCGGACCGT	GGATC M phlei	
2708	ACTCTGA	TIGTATTG	ACATCTAAC	CTCGAACCET	ATATC M.leprae	
1910					M.gastri	
2412	ACTCTGA	TCGTATTG	ACACCTAAC	GTCGAACCCT	GAATC M.kansasii	
4862	ACTCTGA	TCGTATTG	GCCTCTAAC	CTCGGACCGT	ATATC M.smegmatis	
						
					· · · · · ·	
		2490	2500	2510	2520	
3425	GGGTTTA	GGGACAGT	CCTGGCGG	TAGTTTAACT	GGGGC M.tuberculosis	3
364	GGGTTTA	GGGACAGT	CCTGGCGGG	TAGTTTAACT	GGGGC M.bovis	•
2724	GGGTTCA	GGACAGT	CCTGGCGGG	TAGTTTAACT	GGGGC M.avium	
472	GGGTTCA	CGGACAGT6	CCTGGCGGG	TAGTTTAACT	GGGGC M.intracellular	ire
2724	GGGTTCA	GGACAGT	CCTGGCGGG	TAGTTTAACT	GGGGC M.paratuberc.	
5811	QGGTTQA	GGGACAGTO	CCTGGTGGG	TAGTTTAACT	GGGGC M.phlei	
2748	GGTTTA	GGGACAGTG	CCTGGCGGG	TAGTTTAACT	GGGGC M.leprae	
1910		-			M.gastri	
2452	GGGTTCA	D GGACAGTO	CCTGGCGGG	TAGTTTAACT	GGGGC M.kansasii	
4902	QGGTTQA	GGGACAGT	CCTGGIIGGG	TAGTTTAACT	GGGGC M.smegmatis	
				_		
		2930	2940	2950	2960	
3864	AGTACGA	GAGGACCG	GACGGACGA	ACCTCTGGT	CACCA M.tuberculosis	5
3163	AGTACGA	GAGGACCG	GACGGACGA	ACCTCTGGT	ATACCA M.avium	
3163	AGTACGA	GAGGACCG	GACGGACGA	ACCTCTGGT	ATACCA M. paratuberc.	
3256	AGTACGA	GAGGACCG	GACGGACGA	ACCTCTGGT	ATACCA M.phlei	
3187	AGTACGA	GAGGACCG	GACGGACGA	ACCTCTGGT	ATACCA M.leprae	
1910					M.gastri	
2891	AGTACGA	GAGGACCG	GACGGACGA	ACCTCTAGT	GCACCA M.kansasii	
5342	AGTACGA	GAGGACCG(GACGGACGA	ACCTCTGGT	ATACCA M.smegmatis	

Figure 11

		2970	2980	2990	3000
04	GTTGT	CCCCCCAGG	GGCACCGCTG	FATAGCCACG	TTCGGT M.tuberculos
03	GTTGT	CCCACCAGG	GGCACEGCTG	SATAGCCACG	TTCGGA M.avium
03	GTTGT	CCCACCAGG	GGCACGGCTG	SATAGCCACG'	TTCGGA M.paratubero
96	GTTGT	CCCACCAGG	GGCACCGCTG	SATAGCCACG	TTCGGA M.phlei
27	GTTGT	CTICACCAGG	GGCACCGCTG	ATAGCCACG'	TTCGGA M.leprae
10	1				M.gastri
31	GTTGT	CCCACCAGG	GGCACCGCTG	SATAGOTACG	TTCGGA M. kansasii
02	כיייייכייי	حضطفاحته بدور	GGCACEGCTG		
82				SATAGCCACG!	TTCGGM M.smegmatis
02		3010	3020	3030	TTCGGM M.smegmatis
44	CAGGAT	3010	3020	3030	3040
44 43	CAGGAT	3010 PAACCGCTGA	3020 AAAGCATCTAA	3030 AGCGGGAAACO	3040 CTTCTC M.tuberculos
44 43 43	CAGGAT CAGGAT	3010 PAACCGCTGA PAACCGCTGA	3020 NAAGCATCTAA NAAGCATCTAA	3030 AGCGGGAAACO AGCGGGAAACO AGCGGGAAACO	3040 CTTCTC M.tuberculos CTTCTC M.avium
44 43 43	CAGGAT CAGGAT CAGGAT CAGGAT	3010 FAACCGCTGA FAACCGCTGA FAACCGCTGA	3020 NAAGCATCTAA NAAGCATCTAA NAAGCATCTAA	3030 AGCGGGAAACO AGCGGGAAACO AGCGGGAAACO	3040 CTTCTC M.tuberculos CTTCTC M.avium CTTCTC M.paratuberc
44 43 43 36 57	CAGGAT CAGGAT CAGGAT CAGGAT	3010 FAACCGCTGA FAACCGCTGA FAACCGCTGA	3020 NAAGCATCTAA NAAGCATCTAA	3030 AGCGGGAAACO AGCGGGAAACO AGCGGGAAACO	3040 CTTCTC M.tuberculos CTTCTC M.avium CTTCTC M.paratuberc
44 43 43 67	CAGGAT CAGGAT CAGGAT CAGGAT	3010 PAACCGCTGA PAACCGCTGA PAACCGCTGA PAACCGCTGA	3020 NAAGCATCTAA NAAGCATCTAA NAAGCATCTAA NAAGCATCTAA	3030 AGCGGGAAACO AGCGGGAAACO AGCGGGAAACO AGCGGGAAACO	3040 CTTCTC M.tuberculos CTTCTC M.avium CTTCTC M.paratuberc CTCTC M.phlei CTTCTC M.leprae M.qastri
	CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT	3010 PAACCECTER PAACCECTER PAACCECTER PAACCECTER PAACCECTER	3020 NAAGCATCTAA NAAGCATCTAA NAAGCATCTAA	3030 AGCGGGAAACO AGCGGGAAACO AGCGGGAAACO AGCGGGAAACO	3040 CTTCTC M.tuberculos CTTCTC M.avium CTTCTC M.paratuberc CTCTC M.phlei CTTCTC M.leprae M.gastri CTTCTC M.kansasii

	3090	3100	3110	312	0 _
4023		TTCAATAGGT	CAGACCTGGA	AGCT	M.tuberculosis
609	CCCGC-AGAACACGGG!	TTCAATAGG'	CAGACCTGGA	AGCT	M.bovis
3322		ATTGATAGG	CAGACCTGGA	AGCT	M.avium
677	CCCGC-AGACCACGGG!	PTCGATAGG	CAGACCTGGA	AGCT	M.intracellulare
3322	CCCGC-AGATICACGGG	attigataggk	CAGACCTGGA	AGCT	M.paratuberc
3415	CCCGC-AGADCACGGG	TCGATAGAC	CAGACCTGOA	GGCA	M.phlei
3309					M.leprae
1910		•			M.gastri
3050	CCCGC-AGAACACGGG	rtc <mark>c</mark> atagg(CAGACCTGGA	AGCT	M.kansasii
5501	CCCGC-AGAGCACGGG	TEGATAGAC	CAGACCTGGA	AGC	M.smegmatis

Figure 1J

WO 98/15648

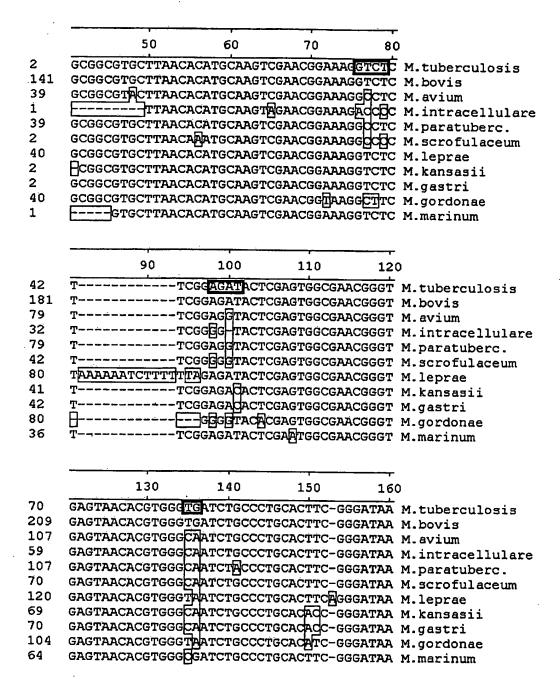


Figure 2A

					
	_	.70	180	190	200
•	GCCTGGGAA	ACTGGGTC	TAATACCGGF	TAGGACCACG	GGA M. tuberculosis
3	GCCTGGGAA	ACTGGGTC'	TAATACCGGA	TAGGACCACG	GGA M.bovis
5	GCCTGGGAA	ACTGGGTC'	TAATACCGGA	TAGGACCTICA	AGA M.avium
	GCCTGGGAA	ACTGGGTC	TAATACCGGA	TAGGACOTT	AGG M.intracellulare
	GCCTGGGAA	ACTGGGTC	TAATACCGG	TAGGACCTCA	AGA M.paratuberc.
	GCCTGGGAA	ACTGGGTC'	TAATACCGG	LA DECTONO CAL	TGG M.scrofulaceum
	GCTTGGGAA	ACTGGGTC	TAATACCGG	TAGGACTION	AGG M.leprae
	GCCTGGGAA	ACTEGETC	TAATACCGG	TAGGACTICA	TGG M.kansasii
	GCCTGGGAA	ACTEGETC'	TABTACCGGA	TAGGACCACI	TGG M.gastri
	GCCTGGGAA	ברייה מכיירי ברייה מכיירי	TANTACCEDA	TREGRECACI	GGA M.gordonae
	CCCTCCCAA	ACTEGETO:	177177CCCM	HOWOOMCOACH	GGA M.marinum
	GCCIGGGAA	ACIGGGIC.	IMMIACCOGA	TAGGACCACG	GGA M.marinum
					
		10	220	230	240
	TGCATGTCT	TGTGGTGG	AAAGCGCTTT	AGCGGTGTGG	GAT M.tuberculosis
	TGCATGTCT	TGTGGTGG?	AAAGCGCTTT	AGCGGTGTGG	GAT M.bovis
	DGCATGTCT	TOTEGTEG	AAAGC-TTTTT	-ACGGTGTGG	GAT M avium
	GCATGTCT	TTAGGTGG	AAAGO-TTT	ECGGTGTGG	GAT M.intracellulare AT M.paratuberc.
	GCATGTCT	Totegrees	AAAGC-TITT	GCGGTGTAG	AAT M. paratuberc
	CGCATGCCT	TGTGGTGG	AAAGCTTT	посестется	GAT M.scrofulaceum
	GCATGTCT	TGTGGTGG	AAAGC-TTTT	песестестве	GAT M.leprae
	GCATGGCT'	TGTGGTGG	AAGCTTT	TCCCCTCTCC	CAT M kengagii
	CGCATGCCT	TGTGGTGG	AAAGC - TTTT	MCCGGTGTGG	GAT M.kansasii GAT M.gastri
	CACATGTCC	Targerge	77760-14171	- GCGGTGTGG	GAT M.gordonae
	TUCATETCO	1 <u>6</u> 1001007	AAGE CTTT	GCGG1G1GG	GAT M.gordonae GAT M.marinum
	-E011010E	101001002		Tiacaa ta ta a	GAT M.marinum
		,			
			260	270	280
	GA GCCGCGG	GCCTATCAG	CTTGTTGGT	GGGTGACGG	CCT M.tuberculosis
	GWCCCCCCC	GCCTATCAG	CTTGTTGGT	GGGGTGACGG	CCT M.bovis
	elecccece	SCCTATCA	CTTGTTGGT	GGGGTGACGG	CCT M.avium
	GGCCCGCG	GCCTATCA G	CTTGTTGGT	gegetgafige(CCT M.intracellulare
	GGGCCCGCG	GCCTATCAG	CTTGTTGGT	GGGGTGACGG	CCT M.paratuberc.
	GGCCCGCG	GCCTATCAG	CTAGTTGGT	ggggtgafigg(CCT M.scrofulaceum
	GEGCCCGCG	SCCTATCAG	CTAATTAGT	gggt¤acgg(CCT M.leprae
	GGGCCCGCG	SCCTATCAG	CTTGTTGGT	GGGTGACGG	CCT M.kansasii
	GGGCCCGCG	SCCTATCAG	CTTGTTGGT	GGGTGACGG	CCT M.gastri
	GE-CCCCCC	GCCTATCAG	CTTGTTGGT	SGGTGAFIGG	CCT M.gordonae
	edecccece	SCCTATCAG	СТТСТТССТ	ടലഭേഷമിച്ചാവര്	CCT M.marinum

Figure 2B

	450	460	470	480	
AAACCT	CTTTCACCA	TCGACGAAGG	TCCGGGTTCT	CTCGG	M.tuberculos
AAACCI	CTTTCACCA	TCGACGAAGG	TCCGGGTTCT		M.bovis
AAACC	CTTTCACCA	TCGACGAAGG	TCCGGGTTAT		M.avium
AAACCT	CTTTCACCA	TCGACGAAGG	TCCGGGTTTT		M.intracellu
			TCCGGGTTTT		M.paratubero
	CTTTCACCA				M.scrofulace
			TCIGGGAATT	CTCGG	M.leprae
AAACCI	CTTTCACCA	TCGACGAAGG	TCCGGGTTCT		M.kansasii
			TCCGGGTTCT		
AAACCI	CTTTCACCA	ICGACGAAGG	TCCGGGTTITT	CTCGG	M.gordonae
AAACCT	CTTTCACCA	rcgacgaagg	TICGGGTTIT	CTCGG	M.marinum

	1130	1140	1150	1160	
1069	TCTCATGTTGCCAGC	ACGTAATGGT	GGGGACTCGT	GAGAG M	1.tuberculosis
1208	TCTCATGTTGCCAGC	ACGTAATGGT	TGGGGACTCGT	GAGAG M	1.bovis
1104	TCTCATGTTGCCAGC	GGGTAATGC	GGGGACTCGT	GAGAG M	1.avium
1056	TCTCATGTTGCCAGC	GGGTAATGC	GGGGACTCGT	GAGAG M	1.intracellulare
1098	TCTCATGTTGCCAGC	gggtaatgca	GGGGACTCGT	GAGAG M	.paratuberc.
1064	TCTCATGTTGCCAGC	GGGTAATGC	GGGGACTCGT	GAGAG M	.scrofulaceum
1119	TCTCATGTTGCCAGC	ACGTAATGGI	GGGGACTCGT	GAGAG M	ileprae
1066	TCTCATGTTGCCAGC	GGGTAATGCC	GGGGACTCGT	GAGAG M	1.kansasii
1067	TCTCATGTTGCCAGC	GGGTAATGCC	GGGGACTCGT	GAGAG M	1.gastri
1100	TCTCATGTTGCCAGC	GGGTAATGCC	GGGGACTCGT	GAGAG M	1.gordonae
1061	TCTCATGTTGCCAGC	ACGTAATGGT	GGGGACTCGT	GAGAG M	1.marinum

	1250	1260	1270	128
1189	CAATGGCCGGTACAA	AGGGCTGCGA	TGCCGCGAGG	TTAAG
1328	CAATGGCCGGTACAA	agggctgcga	TGCCGCGAGG	TTAAG
1224	CAATGGCCGGTACAA	AGGGCTGCGA	TGCCGTAAGG	TTAAG
1176	CAATGGCCGGTACAA	agggctgcga	TGCCGCAAGG	TTAAG
1218	CAATGGCCGGTACAA	agggctgcga	.TGCCGTAAGG	TTAAG
1184	CAATGGCCGGTACAA	agggctgcga	TGCCGCAAGG	TTAAG
1239	CAATGGCCGGTACAA	agggctgcga	.TGCCGCAAGG	TTAAG
	CAATGGCCGGTACAA			
	CAATGGCCGGTACAA			
	CAATGGCCGGTACAA			
1181	CAATGGCCGGTACAA	ひにたにしかにんにか	TRACCARAGA	מ ממוחים

Figure 2C

tuberculosis	1320	1210		•		
bovis					1290	
bovis	CGGGGTCT M.	AGTTCGGAT	TCTC	AAGCCGGT	GAATCCTTA-A	1229
	CGGGGGTCT M	AGTTCGGAT	TCTC	AAGCCGGT	GAATCCTTA-A	1368
Avium	REGEGETET M.	AGTTCGGAT	ACTO	aagccgga	GAATCCTTTTA	1264
intracellulare	TIGGGGTCT M	AGTTCGGAT	TCTC	AAGCCGGT	GAATCCTTTTA	1216
paratuberc.	TIGGGGTCT M	AGTTCGGAT	ACTO	AAGCCGGA	GAATCCTTTTA	1258
scrofulaceum	CGGGGTCT M	AGTTCGGAT	TCTC	AAGCCGGT	GAATCCTTTTA	1224
lenrae	CGGGGGTCT M	AGTTCGGAT	TCTC	AAGCCGGT	GAATCCTTTTA	1279
kangasii	CGGGGTCT M	AGTTCGGAT	TCTC	AAGCCGGT	GAATCCTTTTA	1226
aestri	CGGGGGTCT M	GTTCGGAT	TCTC	AAGCCGGT	GAATCCTTTTA	1227
gascri	CGGGGGTCT M	GTTCGGAT	TCTC	AAGCCGGT	GAATCCTTITA	1260
marinim	CGGGGGTCT M	GTTCGGAT	TCTC	AAGCCGGT	GAATCCTTT	1221
ar IIIam						
		· · · · · · · · · · · · · · · · · · ·		· · ·		
	1360					
tuberculosis	TAATCGCA M.	GTCGCTAG	TCGG	GIGAAGT	CAACTCGACCC	1268
bovis	TAATCGCA M.	GTCGCTAG	TCGG	CGTGAAGT	CAACTCGACCC	1407
avium	TAATCGCA M.	GTCGCTAG	TCGG	ATGAAGT	CAACTCGACCC	1304
intracellulare	TAATCGCA M.	GTCGCTAG	TCGG	CATGAAGT	CAACTCGACCC	1256
paratuberc.	TAATCGCA M.	GTCGCTAG	TCGG	AATGAAGT	CAACTAGACCC	1298
,	TAATCGCA M.	GTCGCTAG	TCGG	CGTGAAGT	CAACTCGACCC	1264
scrofulaceum	ጥልክጥሮፎሮክ M	GTCGCTAG	TCGG	CGTGAAGT	CAACTCGACCC	1319
scrofulaceum leprae		-				
leprae	TAATCGCA M.	GTCGCTAG	rcgg	CGTGAAGT	CAACTCGACCC	1266
leprae kansasii	TAATCGCA M.	GTCGCTAG GTCGCTAG	rcgg rcgg	CGTGAAGT CGTGAAGT	Caactcgaccc Caactcgaccc	1266
leprae kansasii gastri	TAATCGCA M. TAATCGCA M.	GTCGCTAG	TCGG	CGTGAAGT	Caactcgaccc Caactcgaccc Caactcgaccc	1266 1267
av in pa	TAATCGCA M. TAATCGCA M. TAATCGCA M. TAATCGCA M.	AGTCGCTAG AGTCGCTAG AGTCGCTAG AGTCGCTAG	TCGG TCGG TCGG TCGG	ATGAAGT ATGAAGT AATGAAGT CGTGAAGT	CAACTCGACCC CAACTCGACCC CAACTAGACCC CAACTCGACCC	1304 1256 1298 1264

Figure 2D

	50	60	70	80
128	TTCCGAACCCGGAA	GCTAAGCCTGC	CAGCGCCGAT	GATAC M.tuber
39 41	TGCCGAACCCGGAA	GCTAAGCCTGC	CAGCGCCAAI	GATAC M.phlei
3559 5743	TACCGAACCCGGAA	GCTAAGCCTGF	CAGCGCCGAT	GATAC M lenge
J / 43	TCCGAACCCGGAA	GCIMAGCCIGC	CAGCECCGAI	GATAC M. Smegm
	90	100	110	120
168	TGCCCCTCCGEG	TGGAAAAGT	AGGACACCGC	CGAAC M.tuber
79	TGCCCCTCCGGGG-	TGGAAAAGT	AGGGCACCG C	CGAAC M.bovis
81	TGCCCTCACDGGG-	TGGAAAAGT	AGGACACCGC	CGAAC M.phlei
3599	TGCCCATTCGGG	TGGAAAAGT	AGGACACTIGO	CGAAC M.lepra
5782	TACCCITY-CGGGG-	TGGAAAAGT	aggacaccc	CGAAC M.smegm

Figure 3

•	90	100	110	120
382	GGGAGCTGTCAACCGA	GCATTGATC	CGAGGATTTC	CGAAT M. avium
382	GGGAGCTGTCAACCGA	GCATTGATO	CGAGGATTTC	CGAAT M. paratuberc.
L053	GGGAGCTGTCAACCGA	GCGTGGATC	CGAGGATTTC	CGAAT M. tuberculosi
167	GGGAGCTGTCAACCGA	GCGTGGATC	CGAGGATTTC	CGAAT M.phlei
92	GGGAGCTGTCAACCGA	GCGTGGATC	CGAGGATTTC	CGAAT M.leprae
67	GGGAGCTGTCAACCGA	GCGTGGATC	CGAGGATTTC	CGAAT M.gastri
10	GGGAGCTGTCAACCGA	GCGTGGATC	CGAGGATTTC	CGAAT M kansasii
548	GGGAGCTGTCAACCGA	GCGTTGATC	CGAGGATGTC	CGAAT M.smegmatis

	170	180	190	200	ט
462	GAATATATAGG	TGCG-GGAG	TAACGCGGG	AAGTGAAA	M.avium
462	GAATATATAGG	STGCG-GGAGG	TAACGCGGG	SAAGTGAAA	M. paratubero
1133	GAATATATAGG	FTGCG-GGAGG	GAACGCGGG	SAAGTGAAA	M. tuberculosis
547	GAATATATAGG	стде-себсе	GAACGCGGG	BAAGTGAAA	M.phlei
472	GAATATATAGG	TICG-GGAGG	BAACGCGGG	FAAGTGAAA	M.leprae
247	GAATATATAGG	TGCG-GGAGG	GAACGCGGG	Baagtgaaa	M.gastri
190	GAATATATAGG	TGCG-GGAGG	GAACGCGGG	FAAGTGAAA	M.kansasii
2628	GAATATATAGG	CITCII—GGGGG	BAACGCGGGG	FAAGTGAAA	M.smegmatis

		250 .	260	270	280
41	-GTCAG	TAGTGGCGA	GCGAAC-CGG	AACA-GGCTA	AACCG M.avium
41	-GTCAGT	PAGTGGCGA	GCGAAC-CGG	BAACA-GGCTA	AACCG M. paratuher
212	-GCAAG	PAGTGGCGA	GCGAACGCG	BAACA-GGCTA	AACCG M. tuberculo
6	-GTGAG	PAGTGGCGA	GCGAA-AGGG	SAGGATIGGCTA	WAACCG M.phlei
1	-GCAAG1	PAGTGGCGA	GCGAACGTGG	BARIAIGGCTA	AAACCG M.leprae
б	-GTCAG1	TAGTGGCGA	GCGAACGCGG	BAACATGGCTA	AAACCG M.gastri
59	-GTAAG1	ragtggcga	GCGAACGCGG	BAACATGGCTA	AACCG M.kansasii
706	GETGAG	ragtggcga	GCGAACACG	PAGGATGGCTA	AACHG M.smegmatis

Figure 4A

	2	90	300	310	320
578	CATG-CATG	GACAACCG	GTAGGGGTT	TGTGTGCGG	GGT M.avium
578	CATG-CATG	GACAACCG	GTAGGGGTTG	TGTGTGCGG	GGT M.paratuberc.
1250	CAGG-CATG	GTAACCG	GTAGGGGTTG	тстстсссс	GGT M.tuberculosis
664	COTG-CATG	TGATACCO	зетосестте	rerereceef	GT M.phlei
590	CACA-CATG	TCTBACTE	₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽	STETETECEC	GT M.leprae
365	CACG-CATG	GGTGACCG	₹GT⊅GGGGTT0	בייפייפייפרפפ <u>ו</u> .	GGT M.gastri
308	CACG-CATG	GETTE ACCE	36TA6666116	776767676767676 776767676767676	GGT M.kansasii
2745	Marcheare		COTACCCCTTC	TG1G1GCGG(GGT M.kansasii GGT M.smegmatis
2710	FlytoBoxto	IGNINCOG	3G IAGGGG ITC	3161616066	GT M.Smegmatis
			· · · · · · · · · · · · · · · · · · ·		
	3	30	340	350	360
617	mamaadama				- T - T
617	TGTGGGATT	GATATG PC	rcago <u>tetac</u>	CTGGCTGAGG	-GG M.avium
617	TGTGGGATT	'GATATGTC'	rcagc <u>t</u> ctacc	TGGCTGAGG-	-GG M.paratuberc.
1289	TGTGGGAG-	GATATGTC	CAGCGCTACC	PGGCTGAGA	-GG M.tuberculosis
703	TGTGGGGCC	rereret	-CATICGTCCGC	DGGCGATGG	CAG M.phlei
629	TGTGGGATT	GGTATGTC	PCAMCTCTACC	TGGTTGAGG-	-GG M.leprae
404	TGTGGGATC	GATACGTC:	rcagetetace	CGGCTGAGG-	-GG M.gastri
347	TGTGGGATC	GATACGTC	PCAGCTCTACC	GGCTGAGG-	-GG M.kansasii
2785	TGTGGGACC	TATOTITO	CGCCTCTACC	TGGCTGFGAG	GG M.smegmatis
	·				
			,		
	. 3	70	380	390	400
656	TAGTCAGAA	AGTGTCGTC	GUTACCGGAA	GECCCECC	AC M.avium
656	TAGTCAGAA	AGTGTGGTC	GTTAGCGGAP	CMCCCCMCCC	
1327	HAGTCAGAA	ACTOTOGIC ACTOTOGIC	GTTAGCGGAP	CECCEC	
742	MY CUMPY DAY	VGIGICGIO	GTTAGCGGAP GTTAGGTGAP	CONCOCCO CO	
668	TYG TEMTIMA	ACCACION	TAGETEAN	GTGGCCTGGG	, , •
443	Promonos -	AGTGGCGTG	GTTAGCGGAA	MALGGCCTGGG	1 1
386	CAGTCAGAA	AGTGTCGT	GTTAACGGAA	GTGGCCTGGG	FAT M.gastri
	PAGTCAGAA	AGTGTCGT0	GTTANCGGAA	<i>l</i> GTGGCCTGGG	SAT M.kansasii
2823	<u> </u>	AMTGTEGTO	GTTAGCGGAA	ATGGCTTGGG	FAT M.smegmatis

Figure 4B

867

18/31

		•	•			
		410	420	430	440)
696	GGCCCGC	CGTAGACG	GTGAGAGCC	CGGTACGCGAZ	A-ACC	M.avium
696	GGCCCGC	CGTAGACG	GTGAGAGCC	CGGTACGCGA	AA-ACC	M. paratuberc
1367	GGIIGIIGC	CGTAGACG	GTGAGAGCC(CGGTACGCGA	AA-ACC	M.tuberculosis
782	eelichec	CGTAGTGG	GTGAGAGCC	CGTAAC CGA	A-ACA	M.phlei
708	eecciec	CGTAGACG	GTGAGAGCC	AGTACGCGAP	VA-GCC	M.leprae
483	ediciec	CGTAGACG	GTGAGAGCC	CGGTACGTGAA	A-ACC	M.gastri
426	eeliciec	CGTAGACG	GTGAGAGCC	CGGTACGTGAP	A-ACC	M.kansasii
2863	eecciroc	CGTAGACG	GTGAGAGCC	CGGTACGIIGAA	IA-ACC	M. smegmatis
		,			——	
		450	460	470	480	1
735	CGGCACC	TGCCTTAT	ATCAACATC	GAGTAGCAGC	CCCCC	M. aasia
735	CGGCACC	TGCCTTAT	ATCAACACC	ODACORTOADC CORTOATORO	GGGCC	M.paratuberc.
1406	CGGCACC	TGCCTAGT	ATCAAFTICCC	GRETRECREC	GGGCC	M.tuberculosis
820	TIGET GCC	GCTGTCA	CAGGTCCC	GAGTAGCAGC	GGGCC	M. Cubelculosis
747	TGGCACC	TGCCTTGT	ATCAATTCCC	GAGTAGCAGC	GGGCC	M.lenrae
522	CGGCACC'	TGCCTTGT	ATCAATTCCC	GAGTAGCAGC	GGGCC	M.gastri
465	CGGCACC'	TGCCTTGT	ATCAATTCCC	GAGTAGCAGC	GGGCC	M.kansasii
2902	CGACGTC	TGICTTGAT	regretrecc	GAGTAGCAGC	GGGCC	M.smegmatis
						
	•					
			÷ =			**
		· · · · · · · · · · · · · · · · · · ·				
		570	580	590	600	
855	GAGGGAA'	FGGTGAAA	AGTACCCCGG	GAGGG-AGTG	ATAGA	M.avium
855	GAGGGAA	IGGTGAAA	AGTACCCCGG	GAGGG-AGTG	AAATA	M.paratuberc.
1526	GAGGGAA	rggtgaaa <i>i</i>	GTACCCCGG	GAGGGGAGTG	AAAGA	M.tuberculosis
937	GAGGGAAT	r GTGAAA	GTACCCCGG	GAGGG-AGTG	AAAGA	M.phlei
867						

Figure 4C

GAGGGAATGGTGAAAAGTACCCCGGGAGGGGAGTGAAATA M.leprae

GAGGGAATGGTGAAAAGTACCCCGGGAGGGAGTGAAAGA M.gastri
S85 GAGGGAATGGTGAAAAGTACCCCGGGAGGGAGTGAAAGA M.kansasii
3022 GAGGGAATGGTGAAAAGTACCCCGGGAGGGGAGTGAAAGA M.smegmatis

		·	·	
	610	620	630	640
894	GTACCTGAAACC	TGTGCCTACAAT	CCGTCAGAG	CCTCCT M.avium
894	GTACCTGAAACC	TGTGCCTACAAT	CCGTCAGAG	CCTCCT M.paratuberc.
1566	GTACCTGAAACC	STGTGCCTACAAT	CCGTCAGAG	CCTCCT M.tuberculosi
976	GTACCTGAAACC	TGTGCCTACAAT	CCGTCAAAG	CCCTCT M.phlei
907	GTACCTGAAACC	TGTGCCTACAAT	CCGTCAGAG	CCTCTT M.leprae
682	GTACCTGAAACC	STGTGCCTACAAT	CCGTCAGAG	CCCCT M.gastri
625	GTACCTGAAACC	STGTGCCTACAAT	CCGTCAGAG	CCCTTT M.kansasii
3062		TGGGGTTACAAT	CCGTCAGAG	CCCTCG M.smegmatis
				oolog smegmatts
*				
	650			
	650	660	670	680
934	C	GTGGGGTGA	TGGCGTGCC	TTTTGA M.avium
934	C	GTGGGGTGA	TGGCGTGCC	TTTTGA M. paratuberc
1606	TTTCCTCTCCGG	GGAGGGTGGTGA	TGGCGTGCC	TTTTGA M.tuberculosi
1016	Cri	GTAGTGGGGTGA	TGGCGTGCC	TTTTGA M.phlei
947	T	GTGGGGTGA	TGGCGTGCT	TTTTGA M.leprae
722	<u> </u>	GTGGGGTGA	тевсетвеё	TTTTGA M.gastri
665	Č	GTGGGGTGA	TGGCGTGCC	TTTTGA M.kansasii
3102	ACGTGT	GTGGGGTGA	TGGCGTGCC	TTTTGA M.smegmatis
	·			, , , , , , , , , , , , , , , , , , ,
	690	700	710	
				1 - 0
959	AGAATGAGCCTGC	SAGTCAGGGA <u>CAC</u>	GTCGCGAGG	TTAAC M.avium
23	AGAATGAGCCTGC	SAGTCAGGGACAC	GTCGCGAGG	TTAAC M.intracellula:
959	AGAATGAGCCTGC	Gagtcagggacac	GTCGCGAGG	TTAAC M.paratuberc.
				TTAAC M.tuberculosis
4	AGAATGAGCCTGC	Sagtcagggaca <mark>i</mark>	GTCGCAAGG	TTAAC M.bovis
1046	AGAATGAGCCTGC	Sagtcagggaca _l i	GTCGCGAGG	TTAAC M.phlei
972	AGAATGAGCCTGC	SAGTCAGGGACAT	GTCGCGAGG	TTAAC M.leprae
747	AGAATGAGCCTGC	SAGTCAGGGACA <mark>T</mark>	GTCGCGAGG	TTAAC M.gastri
690				TTAAC M.kansasii
3132	AGAATGAGCCTGC	SAGTCAGGGACAL	GTCGCGAGG	TTAAC M.smegmatis

Figure 4D

		770	780	790	800	
1039	CCCATCC	CCTTTEGG		-GTGTAGTGG	CGTGT M	Oszisam
103		CCTTTGGG				intracellulare
1039		CITTTEGG		೨೨ ೯೨ ೩೯೨೪೨–-	CGIGI M.	paratuberc.
1726	CGACCCA	CACGCGCA	TACGCGCGTG	DOTONTOTO DOTONTATION	CGTGT M.	tuberculosis
84	CGACCCA	CACGCGCA	TACGCGCGTG	TGAATAGTGG	CGTGT M	povie
1126	CGTATCC	AACOTGTT		GGTGTAGTGG		
1052	CGHATCA	CGTGTGAG	CGT	-GTGTAGTGG	CGTGT M.	Jenne
827		CGCGTAAG		-GTGTAGTGG	CGTGT M	reptae reptae
770		CCCCCAG		-GTGTAGTGG	CGTGT M	yastri kanadaii
3212		ACACAAGA		-GTGTAGTGG	MGTGT M	smeametic
	u	<u>-2-</u>			<u> </u>	Smegmat 15
					•	
		· · · · · · · · · · · · · · · · · · ·				
		1050	1060	1070	1080	
1307	CAGCCAP	vactccgaa	TGCC <mark>G-TGG</mark>	IG-TAAAAGC	TGGCA M.	.avium
1307	CAGCCAP	<i>l</i> actccgaa	TGCCG-TGG:	rg-ta <u>aaa</u> gco	STGGCA M.	paratuberc.
2005	CAGCCAP	lactccgaa	тессе-тее	IG-TA-AAGC	FTGGCA M.	tuberculosis
1401	CAGCCAP	lactccgaa	TGCCGATAA	TGAAAGIR	FTGGCA M.	.phlei
1323	CAGCCAP	<i>l</i> actccgaa	TGCCG-TGG	TAAAAGC	TGGCA M.	.leprae
1098	CAGCCAA	ACTCCGAA	TGCCG-TGG	ig-talia/gcc	TGGCA M.	.gastri
1041	CAGCCAP	actccgaa	TGCCG-TGG	G-TATA-GCG	TGGCA M.	.kansasii
3486	CAGCCAA	ACTCCGAA	TGCCGGTAAC	GCCAAGAGTG	GGGAA М.	.smegmatis
						•
		•				
		1170	1180	1100		
				1190	1200	
1425	AGTGGAA	AAGGATGT	GTAGTCGCAC	A-GACAACC	GGAGG M.	avium
1425	AGTGGAA	AAGGATGT	GTAGTCGCAG	BA-GACAACCA	AGGAGG M.	paratuberc.
2122	AGTGGGA	AAGGATGT	GCAGTCGCA	A-GACAACC	GGAGG M.	tuberculosis
1519	AGTGGAA	AAGGATGT	GCAGTCGC	PARGACAACCA	GGAGG M.	phlei
1441	AGTGGAA	AAGGATGT	GCAGTCGCA	A-GACAACCA	GGAGG M.	leprae
1215	AGTGGGA	AAGGATGT	GCAGTCGCAG	- BA-GACAACCA	GGAGG M.	gastri
1158	AGTGGGA	AAGGATGT	GCAGTCGCAG	SA-GACAACCA	GGAGG M.	kansasii
3606	AGTGGAA	ÄAGGATGT	GAAGTCGCA	SAAGAAAACCA	GGAGG M.	smegmatis
			_		• • • • • • • • • • • • • • • •	g ± U

Figure 4E

		1250	1260	1270	1280)
1504	CTCACTG	GTCAAGTGA	TATGCGCC	ATAATGTAGC	GGGG	M.avium
1504	CTCACTG	GTCAAGTGA'	TTATGCGCCG	ATAATGTAGC	GGGG	M.paratuberc.
2201	CTCACTG	STCAAGTGA'	TETECCCC	ATAATGTAGC	GGGG	M.tuberculosis
1598	CTCACTG	STCAAGTGA'	TTGTGCGCTG	ATAATGTAGC	GGGG	M.phlei
1520	CTCACTG	GTCAAGTGA'	TTGTGCGCCG	ATAATGTAGC	GGGG	M.leprae
1294	CTCACTG	GTCAAGTGA'	TETECECCE	ATAATGTAGC	GGGG	M.gastri
1237	CTCACTG	STCAAGTGA'	TTGTGCGCCG	ATAATGTAGC	GGGG 1	M.kansasii
3686	TTCACTG	STCAAGTGA'	TTGTGCGCCG	ATAITTGTEGC	GGGG I	M.smegmatis
	_		_	0 0		, , , , , , , , , , , , , , , , , , ,
	<u> </u>			<u> </u>		
		1290	1300	1310	1320	•
1544	CTCAAGC	ACACCGCCG	AAGCCGCGGC	ACATTCATCT	T-TA I	M.avium
1544	CTCAAGC	ACACCGCCG	aagccgcggc	ACATTCATCT		M.paratuberc.
2241	CTCAAGC	ACACCGCCG	AAGCCGCGGC	ACATOCAOCT	TGTH 1	M.tuberculosis
1638	CTCAAGC	ACACCGCCG	aagccgcgc	A-ATCAGCO	TITE I	M.phlei
1560	CTCAAGC	ACACCGCCG	*Yecceceec	ACATTCACCT	TOTA I	M.leprae
			4AGCCGCGAC	- 1		_
1277	CTCAAGC	ACACCGCCG1	AAGCCGCGAC	AJAPOG		M.kansasii
3726	IITCAAGC	ACACCGCCG1	aagccgcgg <u>a</u>	AGOCAACG	TITIG 1	M.smegmatis
				<u> </u>		
		1330	1340	1350	1360	
1583	CGGTGGA	rgtgggtag	ggagcg1 <u>cc</u>	CCCATTCAGC	gaag i	4.avium
1583	CGGTGGAI	rgtgggtag	egagcetcc	CCCATTCAGC	gaag 1	1.paratuberc.
	GGTGGGT	rgtgggtag	eggagcgtcc	CICATICAGO	gaag 1	1.tuberculosis
1676	11CCC1CC1	GTGGGTAG	GGAGCGTCC	TGCATOCGGI	gaag i	1.phlei
1600		GTGGGTAGG	GGAGCGTIC	CICATTCAGC	gaag 1	4.leprae
1367	AGGT			CICATTCAGC		
1310	AGGT			GICATTCAGC		
3764	TT	TGGGTAGG	GGAGCGTCC	tg-atdcggii	GAAG 1	1.smegmatis

Figure 4F

				· · · · · · · · · · · · · · · · · · ·	
		1370	1380	1390	1400
1623	CT-CCG	GGTGACCGG	TGGTGGAGGG	TGGGGGAGTG	AGAAT M.avium
1623	CT-CCG	GGTGATICGG	TGGTGGAGGG	TGGGGGAGTG	AGAAT M.avrum AGAAT M.paratuberc.
2319	CEACCG	GGTGACCGG	TGGTGGAGGG	TGGGGGAGTG	AGAAT M. paratuberc. AGAAT M. tuberculosis
1716	acecce	Астсапссо	TGGTGGAGGG	TCMCCCACTC	AGAAT M.tuberculosis AGAAT M.phlei
1640	CCTCCG	SSCORE AND S	тестселесе: Тестселесе	rececpy cac	AGAAT M.leprae
1402	CCCCCC	GETENCOCC	reeneerech	TGGGGHAGIG	AGAAT M.1eprae AGAAT M.gastri
1345	Calcock	GCTGACCGC	TCCTCCTCCT	TGGGGGAGTG	AGAAT M.gastri
2706	CIGCCG		TIGGIGGAGGA	regeggaete 	AGAAT M.kansasii
3790	A G G C C G	He THT CENE	TGGTGGAGGG	TGEGGGAGTG	AGAAT M.smegmatis
					•
	****	· . , ·			
		1530	1540	1550	1560
1781	CGATGG	ACAACGGGT	TGATATTCCC	GTACCCGTGT	ATGGG M.avium
1781	CGATGG	ACAACGGGT	TGATATTCCC	GTACCCGTGT	ATGGG M. paratuberc
2479	CGATGG	ACAACGGGT	TGATATTCCC	GTACCCGTGT	GTGGG M.tuberculosis
1875	CGATGG	ACAACGGGT	TGATATTCCC	GTACCCGTGT	ATGAG M.phlei
1800	CGATGG	ACAACGGGT	TGATATTCCC	GTACCCGTGT	GTGIG M.leprae
1562	CGATGG	ACAACGGGT	TGATATTCCC	GTACCCGTGT	GTGGG M.gastri
1505	CGATGG	CAACGGGT	ТСАТАТТССС	GTACCCGTGT	GTGGG M.kansasii
3956	CGATGG		TCATATIOCO	GIVCCCGIGI	ATGIG M.smegmatis
0300	CONTOGR	CANOGGG I	IGNINITOOO	GIACCCGTGT	ATGEG M. Smegmatis
					
		1570	1580	1590	1600
1821	CGTCCCT	GATGAATC	A-GCGGTACT	AACCACCCAA	AACCG M.avium
1821	CGTCCCT	GATGAATC	A-GCGGTACT	AACCACCCAA	AACCG M.paratuberc.
2519	CGCCGT	GADGAATC	A-GCGGTACT	BBCCBCCCAR	AACCG M.paratuberc. AACCG M.tuberculosis
	CGTCCCT	רבעבטדעב	TCZ TTCTCC	™™	AACCG M.tuberculosis AACCM M.phlei
1840	CGCCCG	PGDTGDDTC	7-6C6647 Cm	TACCACCCAA)	AACCI M.pniei AACCG M.leprae
1602		ICATOMATO	V-CCCCMVCM	JACCACCCAA	ACCG M. Teprae
1545	CCCCCCC	CANCARTO	R-GCGGTACT/	AACCACCCAA)	AACCG M.gastri
	CCMCCAN	GATGAATC	A-GUGGTACT		AACCG M.kansasii
3330	CGTCCAT	GATGAATC	A~GCGGTACT	AACCA[[CCAA	AACCA M.smegmatis

Figure 4G

					
	1610	1620	1630	164	.0
1860	GAT-CGACCAT-TCC	ССТТСЕСЕС	C-GTGGCGA	TT-CCC	M assissm
1860	GAT-CGACCAT-TCC	CCTTCGGGGG	C-GTGGCGA	<u> </u>	M paratuhana
2558	GAT-CGARCAD-TCC	CCTTCGGGG			M tubopaulogia
1955	GAT-CGATICAD-TCC	E-TrrceeeeE	- Lenebles		W. capetcaloata
1 8 / 9	GAT-CGACCATIATICC	רכייירפפפפפ	・		W lamma
1641	GAT-CGATCAC-TCC	CCTTCGGGGG	B-GTGGNGG		M. Teprae
1584	GAT-CGATCAC-TCC	CCTTCGGGGG	C-CTCCNCC		M. gastri
4035	ACCETGACCECACCE		C_GIGGEG	TOTIOG	M. Kansasii
4033	Pecca Taye checycles		_Terecell	166166	M. Smegmatis
		····			
	1650	1660	1670	168	~
1896	GGCTGCGTGGGACCT	TCGCTGGTAG	TAGTCAAGO	AATEGG	M.avium
1896	GGCTGCGTGGGACCT	TCGCTGGTAG	TAGTCAAGC	AATGGG	M.paratuberc
2594	GGCTGCGTGGGAACT	TCGCTGGTAG	TAGTCAAGC	annagate	M tuberqulogie
1986	GGCTGCGTGGGACCC	G-GTGGGTAG	TAGTCAAGC	ATGGG	M.phlei
1917	GGCTGCGTGGGAACT	TCGITTGGTAG	TAGTCAAGC	STATGGG	M.leprae
1677	GGCTGCGTGGAGCCT	TCGCTGGTAG	TAGTCAAGC	PERTARE	M. castri
1620	GGCTGCGTGGAGCCT	TCGCTGGTAG	TAGTCAAGC	ATGGG	M.kansasii
4071	GGCTGCATGGGACCT	TCGTTGGTAG	TAGTCAAGC	ATGGG	M. smegmatis
	_	_	•	_	3
	•				
					
	1690	1700	1710	172	~
1936	-GTGACGCAGGAAGG	CAGCCGTACC	AGTCAGTGG	FAATA-	M.avium
1936	-GTGACGCAGGAAGGG	CAGCCGTACC	AGTCAGTGG'	FAATA-	M.avium M.paratuberc
1936	-GTGACGCAGGAAGGG	CAGCCGTACC	AGTCAGTGG'	FAATA-	M.avium M.paratuberc
1936 2634 2025	-GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC	AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG'	FAATA- FAATA- FAAGA-	M.avium M.paratuberc. M.tuberculosis M.phlei
1936 2634 2025 1957	-GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC	AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG'	TAATA- TAATA- TAATA- TAATA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae
1936 2634 2025 1957	-GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC	AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG'	TAATA- TAATA- TAATA- TAATA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae
1936 2634 2025 1957 1717	-GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC	AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG	FAATA- FAAGA- FAATA- FAATA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri
1936 2634 2025 1957 1717 1660	-GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC	AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG	FAATA- FAATA- FAATA- FAATA- FAATA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii
1936 2634 2025 1957 1717 1660	-GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC	AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG	FAATA- FAATA- FAATA- FAATA- FAATA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii
1936 2634 2025 1957 1717 1660	-GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG -GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC CAGCCGTACC	AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG'	TAATA- TAAGA- TAAGA- TAATA- TATAA- TATAA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
1936 2634 2025 1957 1717 1660 4111	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC CAGCCGTACC TAGCCGTACC	AGTCAGTGG'AGTGG'AGTGG'AGTGG'AGTGG'AGTGG'AGTGG'AGTGG'AGTGG'AGTG'AGTGG'AGTG'AGTGG'AGTG'A	FAATA- FAATA- FAATA- FAATA- FAATA- FAATA- FAATA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
1936 2634 2025 1957 1717 1660 4111	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-CTGACGCAGGAAGGG-CTGACGCAGGAAGGG-CTGACGCAAGC	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC	AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG' AGTCAGTGG'	FAATA- FAATA- FAATA- FAATA- FAATA- FAATA- FAATA- FAATA-	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
1936 2634 2025 1957 1717 1660 4111	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-CTGGGGCAAGCCCGG-CTGGGGCAAGCCCGG-CTGGGGCAAGCCCG-CTGGGGCAAGCCCG	CAGCCGTACC CAGCCGTACC TAGCCGTACC CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC	AGTCAGTGG'	FAATA- FA	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
1936 2634 2025 1957 1717 1660 4111 1974 1974 2672	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC	AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG	FAATA- FA	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis M.avium M.paratuberc. M.tuberculosis
1936 2634 2025 1957 1717 1660 4111 1974 1974 2672 2063	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-GTGACGCAGGAAGGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG-CTGGGGCAAGCCGG-CCGGGGAAGCCGG-CCGGGGGAAGCCGG-CCGGGGGAAGCCGG-CCGGGGGAAGCCGG-CCGGGGGAAGCCGG-CCGGGGGAAGCCGG-CCGGGGGAAGCCGG-CCGGGGGAAACCCGG-CCGGGGGAAGCCGGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAACCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAACCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAACCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAACCCGG-CCGGGGGAAAACCCGG-CCGGGGGAAAACCCGG-CCGGGGGAAAACCCGG-CCGGGGGAAAACCCGG-CCGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGGAAGCCCGG-CCGGAAGCCCGG-CCGGAAGCCCGG-CCGGGAAGCCCGG-CCGGAAGCCCGG-CCGGAAGCCCGG-CCGGAAGCCCGG-CCGGAAGCCCGGAAGCCCGG-CCGGAAGCCCGGAAGCCCGGAAGCCCGGAAGCCCGGAAGCCCGGAAGCCCGGAAGCCCGGAAGCCCGGAAGCCCGGAAGCCAGAAGCCGGAAGCCAGAAGCCAGAACCCGGAAGCCAGAAGCCAGAAGCAGAACCAGAAGCAAGAACCAGAAACCAGAAACCAGAAACCAGAAACAAACAAAACAAAAACAAAAAA	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC	AGTCAGTGG' AGTAGGCAAI	TAATA- TACGT	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis M.avium M.paratuberc. M.tuberculosis M.phlei
1936 2634 2025 1957 1717 1660 4111 1974 1974 2672 2063 1995	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGCGGTGACGCAAGCCCGGGGAAGCCCGGGGGGAAGCCCGGGGGGAAGCCCGGGGGAAGCCCGGGGGAAGCCCGGGGGAAGCCCGGGGGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC	AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGAG AGTCAGTGAG AGTCAGTCAG	TAATA- TA	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis 0 M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae
1936 2634 2025 1957 1717 1660 4111 1974 1974 2672 2063 1995 1755	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGCCGGGGGGGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGGCAGAGC TAGGCAGAGC TAGGCAGAGC	AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGG AGTCAGTGAA GATAGGCAA GATAGGCAA GATAGGCAA	TAATA- TAATA- TAATA- TAATA- TAATA- TAATA- TAATA- TAATA- TAATA- TACGT TCCGT TCCGT TCCGT	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis O M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri
1936 2634 2025 1957 1717 1660 4111 1974 1974 2672 2063 1995 1755 1698	-GTGACGCAGGAAGGG-GTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGGGGTGACGCAGGAAGCGGTGACGCAAGCCCGGGGAAGCCCGGGGGGAAGCCCGGGGGGAAGCCCGGGGGAAGCCCGGGGGAAGCCCGGGGGAAGCCCGGGGGG	CAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC CAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGTACC TAGCCGAGAGC TAGGCAGAGC TAGGCAGAGC TAGGCAGAGC TAGGCAGAGC	AGTCAGTGG' AGTCAGTAGGCAAI GATAGGCAAI GATAGGCAAI GATAGGCAAI	TAATA- TAATA- TAATA- TAATA- TAATA- TAATA- TAATA- TAATA- TACGT ATCCGT ATCCGT ATCCGT	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis O M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii

Figure 4H

					•
	1810	1820	1830	184	10
2051	CG-AATTCGGTG	ATCCTCTGCTGC	CDDGDDDDGC	ריייריייא –	M eszism
2051	CG-AATTCGGTG	\TCCTCTGCTGC	CARCANANCC		M.paratuberc.
2751	CG-ABTTCGGTG	,TCCTCTCCTCC	CNACARARCC		M. tuberculosis
2141	CG-AATTCGGTGA	7.LCC101010100 7.LCC1010100100	CARGAMARGO	CHCMA-	M. tuberculosis
2074	CG-AATTCGGTA		CDAGAMAAGC	CTCTA	M.pniei
1834	CG-AATTCGGTG	AGCCLCIGCIGC	CAAGAAAAGC	CTCTA-	M.leprae
1777	CG-AATTCGGTG		CAAGAAAAGC	CTCTA-	M.gastri
1220	CG-MAIICGGIGA	TCCTCTGCTGC	CAAGAAAAGC	CTCTA-	M.kansasii
7220	CG-AATTCGGTG/	erccipreciec	CHAGAAAAGC	CTCTA-	M.smegmatis
	1050	40.00		т	
	1850	1860	1870	188	=
2089	GCGAGCACATACA	CGGCCCGTACCC	CCAAACCAAC	ACAGGT	M.avium
2089	GCGAGCACATACA	CIGCCCGTACCC	CCAAACCAAC	ACAGGT	M.paratuberc.
2789	GCGAGCACACACA	CGGCCCGTACC	CCAAACCGAC	CAGGT	M.tuberculosis
2179	GCAAGCECATACA	CGGCCCGTACCC	CCAAACCAAC	CAGGT	M.phlei
2112	GCGAGCACACACA	CGGCCCGTACCC	CCAAACCGACA	CAGGT	M.leprae
1872	GCGAGCACACA	CGGCCCGTACCC	CAAACCGACA	ACAGG	M.gastri
1815	GCGAGCACACACA	CGGCCCGTACCC	CAAACCGACA	CAGGT	M.kansasii
4266	GCGAGGACATACA	CGGCCCGTACCC	CAAACCAACA	CAGGT	M. smeamatis
	. •				···-g
	•				
	1970	1980	1990		•
				200	•
2208	AGGGGCCCCGGAA	<u> </u>			** !
2200		TACLGTGAACAC	CCTTGCGGTG	GGAGC	M.avium
2200	AGGGGGCCCGGAA	TACCGTGAACAC	CCTTGCGGTG	GGAGC	M. paratubero
2908	AGGGGGCCCGGAA AGGGGGACCGGAA	taccgtgaacac taficgtgaacac	CCTTGCGGTG	GGAGC	M.paratuberc.
2908 2298	AGGGGGCCCGGAA AGGGGGACCCACG	TACCGTGAACAC TA∏CGTGAACAC TACCGTGAEGGC	CCTTGCGGTG CCTTGCGGTG MCTTGCGGGG	egagc egagc edagc	M.paratuberc. M.tuberculosis M.phlei
2908 2298 2231	AGGGGGCCCGGAA AGGGGGACCCACG	TACCGTGAACAC TA∏CGTGAACAC TACCGTGAEGGC	CCTTGCGGTG CCTTGCGGTG MCTTGCGGGG	egagc egagc edagc	M.paratuberc. M.tuberculosis M.phlei
2908 2298 2231 1910	AGGGGGCCCGGAA AGGGGGGACCCACG AGGGGGGCCGGAA	TACCGTGAACAC TA[[CGTGAACAC TACCGTGA <u>EGG</u> C TA[[CGTGAACAC	PTOPOSTTONS TOPOSTONS TOPOSTONS	eggagc eggagc eggagc	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri
2908 2298 2231 1910 1934	AGGGGGCCGGAA AGGGGGACCCACG AGGGGGGCCGGAA AGGGGGGACCGGAA	TACCGTGAACAC TA∏CGTGAACAC TACCGTGA <u>EGG</u> C TA∏CGTGAACAC TACCGTGAACAC	CCTTGCGGTGCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTG	GGAGC GGAGC GGAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii
2908 2298 2231 1910 1934	AGGGGGCCGGAA AGGGGGACCCACG AGGGGGGCCGGAA AGGGGGGACCGGAA	TACCGTGAACAC TA∏CGTGAACAC TACCGTGA <u>EGG</u> C TA∏CGTGAACAC TACCGTGAACAC	CCTTGCGGTGCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTG	GGAGC GGAGC GGAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii
2908 2298 2231 1910 1934	AGGGGGCCCGGAA AGGGGGGACCCACG AGGGGGGCCGGAA	TACCGTGAACAC TA∏CGTGAACAC TACCGTGA <u>EGG</u> C TA∏CGTGAACAC TACCGTGAACAC	CCTTGCGGTGCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTG	GGAGC GGAGC GGAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii
2908 2298 2231 1910 1934	AGGGGGCCGGAA AGGGGGACCCACG AGGGGGGCCGGAA AGGGGGGACCGGAA	TACCGTGAACAC TA∏CGTGAACAC TACCGTGA <u>EGG</u> C TA∏CGTGAACAC TACCGTGAACAC	CCTTGCGGTGCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTGCCCTTGCGGTG	GGAGC GGAGC GGAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii
2908 2298 2231 1910 1934	AGGGGGCCCGGAA AGGGGGGACCCACG AGGGGGGCCCGGAA AGGGGGGACCGAA AGGGGGACCCACA	TACCGTGAACAC TA∏CGTGAACAC TA∏CGTGAACAC TA∏CGTGAACAC TACCGTGAACAC TACCGTGAACAC	CCTTECGETE CCTTECGETE CCTTECGETE CCTTECGETE CCTTECGETE	GGAGC GGAGC GGGAGC GGGAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
2908 2298 2231 1910 1934 4385	AGGGGGCCCGGAA AGGGGGGACCCACG AGGGGGGCCCGGAA AGGGGGGACCGGAA AGGGGGACCCACA	TACCGTGAACAC TAIICGTGAACAC TAIICGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC	CCTTGCGGTG CCTTGCGGTG CCTTGCGGTG CCTTGCGGTG	GGAGC GGAGC GGAGC GGAGC CCAAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
2908 2298 2231 1910 1934 4385	AGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCCACA 2010 GGGATTCGGCCGC	TACCGTGAACAC TAIICGTGAACAC TAIICGTGAACAC TACCGTGAACAC TACCGTGAACAC TGGCGTGIAAGC	CCTTGCGGTG CCTTGCGGTG CCTTGCGGTG CCTTGCGGTG CCTTGCGGTG	GGAGC GGAGC GGAGC GGAGC CCAAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
2908 2298 2231 1910 1934 4385 2248 2248	AGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCCACA 2010 GGGATTCGGCCGC	TACCGTGAACAC TAIICGTGAACAC TAIICGTGAACAC TACCGTGAACAC TACCGTGAACAC TEGCGTGIAAGC 2020 AGAAACCAGTG	CCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTTACGGCGCGCTGGTAGCGACTGGTAGCGACTGGTAGCGACTGGTAGCGACTGCCCACTGCCACTGCCGACTGCCGACTGCCGACTGCCGACTGCCGACTGCCGACTGCCCACTGCACTGCCACTGCACTACTACTACTACTACTACTACTACTACTACTACTACT	GGAGC GGAGC GGAGC GGAGC CAAGC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis 0 M.avium M.paratuberc.
2908 2298 2231 1910 1934 4385 2248 2248 2948	AGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGACCCACA 2010 GGGATTCGCCCGC GGGATTCGCCCGC	TACCGTGAACAC TAIICGTGAACAC TAIICGTGAACAC TACCGTGAACAC TACCGTGAACAC TEGCGTGIAAGC 2020 AGAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAAGTGGAAAACCAGTGGAAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAAACCAAGTGGAAAACCAAGTGGAAAACCAAGTGGAAAAACCAAGTGAAAAACCAAGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGGAAAAACCAAGTGGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAAAA	CCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGTGCGACT-GGGAGCGACT-GGGAGCGACT-GGGAGCGACT-	GGAGC GGAGC GGAGC GGAGC CAAGC 204 GTTTA	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis 0 M.avium M.paratuberc. M.tuberculosis
2908 2298 2231 1910 1934 4385 2248 2248 2948 2338	AGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGCAA AGGGGGACCCACA 2010 GGGATTCGGCCGC GGGATCCGCTGCC GGGATCCGCTGCC	TACCGTGAACAC TAIICGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TEGCGTGIAAGC 2020 AGAAACCAGTGI AGAAACCAGTGI AGAAACCAGTGI	CCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCCGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGGT	GGAGC GGAGC GGAGC GGAGC CAAGC CAACC CAAGC CAACC CAAGC CAAGC CAAGC CAACC CAAGC CAACC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis O M.avium M.paratuberc. M.tuberculosis M.phlei
2908 2298 2231 1910 1934 4385 2248 2248 2948 2338 2271	AGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCCACA 2010 GGGATTCGGCCGC	TACCGTGAACAC TAIICGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TEGCGTGIAAGC 2020 AGAAACCAGTGI AGAAACCAGTGI AGAAACCAGTGI	CCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCCGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGGT	GGAGC GGAGC GGAGC GGAGC CAAGC CAACC CAAGC CAACC CAAGC CAAGC CAAGC CAACC CAAGC CAACC	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis O M.avium M.paratuberc. M.tuberculosis M.phlei
2908 2298 2231 1910 1934 4385 2248 2248 2948 2338 2271 1910	AGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGGCCCGGAA AGGGGGACCCACA 2010 GGGATTCGGCCGC GGGATCCGGTCGC GGGATCCGGTCGC GGGATCCGGTCGC	TACCGTGAACAC TAIICGTGAACAC TAIICGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC AGAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAACCAGTGGAAACCAGTGGAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAAACCAGTGGAAAACCAGTGGAAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGGAAAACCAGTGAAAACCAGTGAAAAACCAGTGAAAAACCAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAAAAA	CCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCCGTGCGGTGCGGTGCGGCG	GGAGC GGAGC GGAGC GGAGC CAAGC CAAGC CTTTA GTTTA GTTTA GTTTA	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis O M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri
2908 2298 2231 1910 1934 4385 2248 2248 2948 2338 2271 1910	AGGGGCCCGGAA AGGGGGACCCACA AGGGGGACCGGAA AGGGGGACCCACA AGGGGGACCCACA 2010 GGGATTCGCCGCC GGGATCCGCTGCC GGGATCCGCTGCC GGGATCCGCTGCC GGGATCCGCTCGC GGGATCCGCTCGC GGGATCCGCTCGC	TACCGTGAACAC TAICGTGAACAC TAICGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACAC TACCGTGAACCAGTGAAAACCAGTGAAAACCAGTGAAACCAGTGAAACCAGTGAAAAACCAGTGAAAAACCAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAACCAAGTGAAAAAACCAAGTGAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAACCAAGTGAAAAAAACCAAGTGAAAAAAAA	CCTTGCGGTGCCTTGCGGTGCCTTGCGGTGCGGCGCGCTTGCGGTGCGGCG	GGAGC GGAGC GGAGC GGAGC CCAAGC 204 GTTTA GTTTA GTTTA GTTTA	M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis 0 M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii

Figure 4I

	2130	21	40 2:	150	2160
2367		TAAGGG	rgaagcggag	AATTTAAG	CCC M.avium
2367 3067	CCGTTAACCCG	TAAGGG	rgaagcggag	AATTTAAG	CCC M.paratuberc.
2457	CCGTTAACCCG	U-AAGGG	I'GAAGCGGAG	AATTTAAG	CCC M.tuberculosi
2390	Cherrancech	IIICGGGGG	I GAAGCGGAG FGAAGCGGAG	AATTTAAG	CCC M.phlei CCC M.leprae
1910	010111110000	<u> </u>	. CAROCOAGO	MATTIMAG	M.gastri
2094		-AAGGG1	rgaagcggag	AATTTAAG	CCC M kansasii
4544	CCGTTAACCC	CTTGGGGG	rgaagcggag	AATTTAAG	CCC M.smegmatis
		-		, 	
	2250	22	60 22	270	2280

	2230	2260	2270	2280	
2485	GTAACGACTTCCC	ACTGTCTCAAC	CATAGACTO	GGCGAA M.	avium
2485	GTAACGACTTCCC	VACTGTCTCAAC	CATAGACTC	GGCGAA M.	naratuhero
3185	GTAACGACTTCTC	VACTGTCTCAAC	CATAGACTC	GGCGAA M.	tuberculosis
2577	GTAACGACTTCTC	ACTGTCTCAAC	CATAGACTC	GCGAA M.	phlei
2508	GTAACGACTTCTC	ACTGTCTCAAC	CATAGACTC	GCGAA M.	leprae
1910	·			M.	gastri
2212	GTAACGACTTCTC	ACTGTCTCAAC	CATAGACTC	зесерь м	Kengagii
4663	GTAACGACTTCTCA	LACTGTCTCAAC	∃atagact co	egcgaa m.	smegmatis

	2370	2380	2390	2400
2605	GTTCGGTACGGTT	TGTGTAGGA	TAGGTGGGAGACT	TTGAA M.avium
3305	GTTCGGTACGGTT GTTCGGTACGGTT	TGTGTAGGA TGTGTAGGA	TAGGTGGGAGACT TAGGTGGGAGACT	TTGAA M.paratuberc. GTGAA M.tuberculosis
2697	GOTCGATACGGTT	TGTGTAGGA	TAGGTGGGAGACT	GTGAA M.phlei
2628 1910	GTTCGGTGCGGTT	TGTGTAGGA	TAGGTGGGAGACT	GTGAA M.leprae
	GTTCGGTACGGTT	TGTGTAGGA	TAGGTGGGAGACT	M.gastri GTGAA M.kansasii
4782	GOTCGATACGGTT	TGTGTAGGA ⁽	TAGGTGGGAGACT	GTGAA M.smegmatis

Figure 4J

26/31

	2410	2420	2430	2440	
2645	GCACAFACGCC	GTTTGTGTGG	AGTCGTTGTTG	AAATACC M.	.avium
393	ATACAGACGCC	\GTTTGTATGG	AGTCGTTGTTG	АВАТАСС М	intracellulare
2645	GCACAGACGCC	GTTTGTGTGG:	AGTCGTTGTTG	AAATACC M.	paratubero
3345	ACCTOGACGCCA	GTTGGGGGGG	AGTCGTTGTTG	AAATACC M.	tuberculosis bovis
284	ACCTOGACGCCA	GTTGGGGGGG	AGTCGTTGTT G	AAATACC M.	bovis
2737	GCTICGGACGCC	\GTT C G GTGG	AGTCGTTGTTG	АААТАСС М.	phlei
2668	ACTTOGACGCTP	GTTGGGGTGG:	AGTCGTTGTTG	AAATACC M.	leprae
1910				M.	gastri
2372		GTTGGGGTGG:	AGTCGTTGTTG	АААТАСС М.	kansasii
4822	GCTCADACGCCA	GTGTGGGTGG	agtcgttgttg	алатасс м.	smegmatis
	0.450	0.450			•
•	2450		2470	2480	
2685	ACTCTGATCGTA	TTGGACACCT	AACGTCGAACC	CT-TATC M.	avium
433	ACTCTGATCGT	ATTGGACACCT	AACGTCGAACC	CT-TATC M	intracellulare
2685	ACTCTGATCGT	ATTGGACACCT	aac <u>g</u> tcgaacc	CT-TATC M.	paratuberc.
3385	ACTCTGATCGT	TTGGGCARCT	AACOTCGAACC	СТЕВЪТС М	tuberculosis
324	ACTCTGATCGT	TTGGGCARCT	AACOTCGAACC	$\mathtt{C}\mathtt{T}\mathtt{G}\mathtt{A}\mathtt{T}\mathtt{C}$ M.	bovis
2777	ACTCTGATCGTF ACTCTGATCGTF	TTGGGCCTCT	AACOTCGGACC	GТGGAТС M.	phlei
2708	ACTCTGATTGTA	(TTGAACATICT)	AACDICGAACC	БТАТАТ С М.	leprae
1910				м.	gastri
2412	ACTCTGATCGTA	TTGGACACCT	avcēlcēvycc	ствиатс м.	kansasii
4862	ACTCTGATCGTA	(TTGGGCCTCT)	AACCTCGGACC	СПАТАТС М.	smegmatis
					_
					-
					-
					•
		-			-
	2690	2700	2710	2720	
2924	2690 GGTGTCAF	2700	2710	2720	1.avium
2924 2924	2690 GGTGTCACTCAC GGTGTCACTCAC	2700 ACGGATAAAG	2710 GTACCCGGGG	2720 GATAAQGG M	1.avium 1.paratuberc
2924 2924 3625	2690 GGTGTCACTCAA GGTGTCACTCAA GGTGTCCCTCAA	2700 ACGGATAAAG ACGGATAAAG	2710 GTACCCGGGG GTACCCGGGG	2720 SATAACAG M SATAACAG M	1.avium 1.paratuberc. 1.tuberculosis
2924 2924 3625 3017	2690 GGTGTCACTCAC GGTGTCCCTCAC GGTGTCCCTCAC	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG	2720 FATAACAG M FATAACAG M FATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei
2924 2924 3625 3017 2948	2690 GGTGTCACTCAA GGTGTCACTCAA GGTGTCCCTCAA	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG	2720 GATAACAG M GATAACAG M GATAACAG M GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae
2924 2924 3625 3017 2948 1910	2690 GGTGTCACTCAC GGTGTCCCTCAC GGTGTCCCTCAC GGTGTCCCTCAC GGTGTCCCTCAC	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG	2720 GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri
2924 2924 3625 3017 2948 1910 2652	2690 GGTGTCAF GGTGTCCCTCAF GGTGTCGCTCAF GGTGTCGCTCAF	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG	2720 GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii
2924 2924 3625 3017 2948 1910	2690 GGTGTCACTCAC GGTGTCCCTCAC GGTGTCCCTCAC GGTGTCCCTCAC GGTGTCCCTCAC	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG	2720 GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii
2924 2924 3625 3017 2948 1910 2652	2690 GGTGTCAF GGTGTCCCTCAF GGTGTCGCTCAF GGTGTCGCTCAF	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG	2720 GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii
2924 2924 3625 3017 2948 1910 2652	2690 GGTGTCAF GGTGTCCCTCAF GGTGTCGCTCAF GGTGTCGCTCAF	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG	2720 GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii
2924 2924 3625 3017 2948 1910 2652 5102	2690 GGTGTCALTCAL GGTGTCGCTCAL GGTGTCGCTCAL GGTGTCGCTCAL GGTGTCGCTCAL GGTGTCGCTCAL 2730	2700 ACGGATAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG	2710 GTACCCGGGGGGTACCCGGGGGGTACCCCGGGGGGGGGG	2720 GATAACAG M	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis
2924 2924 3625 3017 2948 1910 2652 5102	2690 GGTGTCAP GGTGTCCCTCAP GGTGTCGCTCAP GGTGTCGCTCAP GGTGTCGCTCAP GGTGTCGCTCAP GGTGTCGCTCAP GGTGTCGCTCAP	2700 ACGGATAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG GTACCCGGGG GTACCCGGGG GTACCCGGGG	2720 GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii 1.smegmatis
2924 2924 3625 3017 2948 1910 2652 5102	2690 GGTGTCACTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GCTGATCTTCCC	2700 ACGGATAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG GTACCCGGGG GTACCCGGGG TACCCGGGG TACCCGGGG	2720 GATAACAG M	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii 1.smegmatis
2924 2924 3625 3017 2948 1910 2652 5102 2964 2964 3665	2690 GGTGTCACTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC 2730 GCTGATCTTCCC GCTGATCTTCCC	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG	2710 GTACCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGTACCCCGGGGTACCCCGGGGTATCGACGGTATCGACGACGACGACGACGACGACGACGACGACACACAC	2720 GATAACAG M GATAAC	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii 1.smegmatis 1.avium 1.paratuberc. 1.tuberculosis
2924 2924 3625 3017 2948 1910 2652 5102 2964 2964 3665 3057	2690 GGTGTCACTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GCTGATCTTCCC GCTGATCTTCCC GCTGATCTTCCC	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG	2710 GTACCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGG	2720 GATAACAG M GATAAC	1.avium 1.paratuberc. 1.tuberculosis 1.phlei 1.leprae 1.gastri 1.kansasii 1.smegmatis 1.avium 1.paratuberc. 1.tuberculosis 1.phlei
2924 2924 3625 3017 2948 1910 2652 5102 2964 2964 3665 3057	2690 GGTGTCACTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC 2730 GCTGATCTTCCC GCTGATCTTCCC	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG	2710 GTACCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGGTACCCCGGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGGTATCGACGGG	2720 GATAACAG M GATAAC	M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae M.gastri M.kansasii M.smegmatis M.avium M.paratuberc. M.tuberculosis M.phlei M.leprae
2924 2924 3625 3017 2948 1910 2652 5102 2964 2964 3665 3057 2988 1910	2690 GGTGTCAP GGTGTCAP GGTGTCGCTCAP GGTGTCGCTCAP GGTGTCGCTCAP GGTGTCGCTCAP GGTGTCGCTCAP GCTGATCTTCCC GCTGATCTTCCC GCTGATCTTCCC GCTGATCTTCCC GCTGATCTTCCC	2700 ACGGATAAAG ACGGATAAAG ACGGATAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATAAAAG ACGGATCCA ACAAGAGTCCA	2710 GTACCCGGGG GTACCCCGGGG GTACCCCGGGG GTACCCCGGGG GTACCCCGGGG TACCCCGGGG TACCCCGGGG TACCCCGGGG TACCCCGGGG TACCCCGGGG TATCGACGGGG TATCGACGGGG TATCGACGGGG TATCGACGGGG	2720 GATAACAG M GATAAC	f.avium f.paratuberc. f.tuberculosis f.phlei f.leprae f.gastri f.kansasii f.smegmatis f.avium f.paratuberc. f.tuberculosis f.phlei f.leprae f.gastri
2924 2924 3625 3017 2948 1910 2652 5102 2964 2964 3665 3057 2988 1910 2692	2690 GGTGTCACTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GGTGTCGCTCAC GCTGATCTTCCC GCTGATCTTCCC GCTGATCTTCCC	2700 ACGGATAAAG ACGGATAAAAG ACGAGAGTCCA CCAAGAGTCCA CCAAGAGTCCA	2710 GTACCCGGGG GTACCCGGGG GTACCCGGGG GTACCCGGGG GTACCCGGGG TACCCCGGGG TATCGACGGG TATCGACGGGG TATCGACGGGG TATCGACGGGG TATCGACGGGG TATCGACGGGG TATCGACGGGGG TATCGACGGGGG TATCGACGGGGGGGGGG	2720 GATAACAG M GATAAC	f.avium f.paratuberc. f.tuberculosis f.phlei f.leprae f.gastri f.kansasii f.smegmatis f.avium f.paratuberc. f.tuberculosis f.phlei f.leprae f.gastri f.kansasii

Figure 4K

					
	2770	2780	2790	280	0
3004	GCACCTCGATG	TCGGCTCGTCG	CATCCTGGGGC	TCGACCA	M avium
3004	GCACCTCGATG	TCGGCTCGTCG	CATCCTGGGGC	TGGAGCA	M. paratuberc.
3705	GCACCTCGATG	TCGGCTCGTCG	CATCCTGGGGC	TGGAGCA	M.tuberculosis
3097	GCACCTCGATG	TCGGCTCGTCG	CATCCTGGGGC	TGGAGCA	M.nhlei
3028	GCACCTCGATG	TCGGCTCGTCG	CATCCTGGGGC	TGAAGCA	M.leprae
1910					M.gastri
2732	GCACCTCGATG	TCGGCTCGTCG	CATCCTGGGGC	TGGAGCA	M. kansasii
5182	GCACCTCGATG	TCGGCTCGTCG	CATCCTGGGGC	TGGAGCA	M.smegmatis
	2810	2820	2830	204	,
		·	-	2840	
3044	GGTCCCAZEG	TTGGGCTGTTC	GCCC-Attara	GCGGCAC	M.avium
3044	GGTCCCAAGGG'	TTGGGCTGTTC	GCCC-ATTAAA	BCGGCAC	M.paratuberc.
3745	GGTCCCAAGGG	TTGGGCTGTTC	GCCC-ATTAAA(GCGGCAC	M.tuberculosis
3137	GGTCCCAAGGG'	TTGGGCTGTTC	GCCC-ATTAAA(SCGGCAC	M.phlei
3068	GGTCCCAAGGG'	TTGGGCTGTTC	GCCC-ATTAAA(_
1910					M.gastri
2772	GGTCCCAAGGG	TTGGGCTGTTC	GCCC-ATTAAA(SCGGCAC	M.kansasii
5222	GGTCCCAAGGG!	TTGGGCTGTTC	GCCCCATTAAA	FCGGCAC	M.smegmatis
	•	•			
	3050	3060	3070	3080	
3283					.avium
638	CAAGATCAGGTT CAAGATCAGGTT	T-CTCACCCTT	TTAGA GGGATA TTAGAGGGATA	AGGCCC M	.intracellulare
638 3283	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT	T-CTCACCQTT TCCCCACCTT-CTCACCTT	TTAGA GGATA TTAGAGGGATA TTAGAGGGATA	AGGCCC M AGGCCC M	.intracellulare
638 3283 3984	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT	TT-CTCACCTT TT-CTCACCCTT TT-CTCACCCTT	TTAGA <mark>EGGATA</mark> TTAGAGGGATA TTAGAGGGATA TTGGTGGGATA	AGGCCC M AGGCCC M AGGCCC M	.intracellulare .paratuberc.
638 3283 3984 570	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT	TTCTCACCTT TTCTCACCTT TTCTCACCTT TTCTCACCCTC	TTAGASGGATA TTAGAGGGATA TTAGAGGGATA TTGGTGGGATA TTGGTGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis
638 3283 3984 570 3376	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT	TTCTCACCTT TTCTCACCTT TTCTCACCTT TTCTCACCCTC	TTAGASGGATA TTAGAGGGATA TTAGAGGGATA TTGGTGGGATA TTGGTGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis
638 3283 3984 570 3376 3307	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT	TTCTCACCTT TTCTCACCTT TTCTCACCTT TTCTCACCCTC	TTAGASGGATA TTAGAGGGATA TTAGAGGGATA TTGGTGGGATA TTGGTGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis
638 3283 3984 570 3376 3307 1910	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT	TTOTOACCOTT TTOTOACCTT TTOTOACCTT TTOTOACCTT TTOTOACCTC TTOTOACCTC	TTAGASGGATA TTAGAGGGATA TTAGAGGGATA TTGGIGGGATA TTGGIGGGATA TAGGAGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M AGGCCC M M	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae
638 3283 3984 570 3376 3307 1910 3011	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAA	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC	TTAGA GGATA TTAGAGGGATA TTAGAGGGATA TTGGIGGGATA TAGGAGGGATA TAGGAGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M AGGCCC M M M M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri
638 3283 3984 570 3376 3307 1910	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC	TTAGA GGATA TTAGAGGGATA TTAGAGGGATA TTGGIGGGATA TAGGAGGGATA TAGGAGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M AGGCCC M M M M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri
638 3283 3984 570 3376 3307 1910 3011	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAA	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC	TTAGA GGATA TTAGAGGGATA TTAGAGGGATA TTGGIGGGATA TAGGAGGGATA TAGGAGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M AGGCCC M M M M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri
638 3283 3984 570 3376 3307 1910 3011	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAA	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC	TTAGA GGATA TTAGAGGGATA TTAGAGGGATA TTGGIGGGATA TAGGAGGGATA TAGGAGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M AGGCCC M M M M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri
638 3283 3984 570 3376 3307 1910 3011	CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAAGATCAGGTT CAA	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC	TTAGA GGATA TTAGAGGGATA TTAGAGGGATA TTGGIGGGATA TAGGAGGGATA TAGGAGGGATA	AGGCCC M AGGCCC M AGGCCC M AGGCCC M AGGCCC M M M M AGGCCC M	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri
638 3283 3984 570 3376 3307 1910 3011 5462	CAAGATCAGGTT	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC	TTAGA GGGATAI TTAGAGGGATAI TTAGAGGGATAI TTGGIGGGATAI TAGGAGGGATAI TTGGIGGGATAI TAGGAGGGATAI TAGGAGGGATAI	AGGCCC MAGGCCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis
638 3283 3984 570 3376 3307 1910 3011 5462	CAAGATCAGGTT	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCAC	TTAGA GGGATAI TTAGAGGGATAI TTAGAGGGATAI TTGGIGGGATAI TAGGAGGGATAI TTGGIGGGATAI TAGGAGGGATAI TAGGAGGGATAI	AGGCCC MAGGCCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis
638 3283 3984 570 3376 3307 1910 3011 5462	CAAGATCAGGTT CCAAGATCAGGTT CCAAGATCAGGTT CCAAGATCAGGTT CCAAGATCAGGTT CCAAGATCAGGTT CCCGC-AGACCA	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCTC TT-CTCACCCTC TT-CTCACCCTC TT-CTCACCCTC	TTAGA GGATA TTAGAGGGATA TTAGAGGGATA TTGGIGGGATA TTGGIGGGATA TAGGAGGGATA TAGGAGGGATA TAGGAGGGATA 3110	AGGCCC MAGGCCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium
638 3283 3984 570 3376 3307 1910 3011 5462 3322 677 3322	CAAGATCAGGTT CCAAGATCAGGTT CCCGCAGACCAGAC	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCCAC TT-CTCACCCTC TT-CTCACCCTC TT-CTCACCCTC TT-CTCACCCTC	TTAGA GGGATA TTAGAGGGATA TTAGAGGGATA TTGGTGGGATA TTGGTGGGGATA TAGGAGGGATA TAGGAGGGATA 3110 GGCAGACCTGG	AGGCCC MAGGCCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium .intracellulare .paratuberc
638 3283 3984 570 3376 3307 1910 3011 5462 3322 677 3322 4023 609	CAAGATCAGGTT CCAGCAGACCAGAC	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC	TTAGA GGGATAL TTAGAGGGATAL TTAGAGGGATAL TTAGAGGGATAL TAGGAGGGATAL TAGGAGGGATAL TAGGAGGGATAL 3110 GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG	AGGCCC MAGGCCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium .intracellulare .paratuberc .tuberculosis .bovis
638 3283 3984 570 3376 3307 1910 3011 5462 3322 677 3322 4023 609	CAAGATCAGGTT CCAGCAGACCAGAC	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC	TTAGA GGGATAL TTAGAGGGATAL TTAGAGGGATAL TTAGAGGGATAL TAGGAGGGATAL TAGGAGGGATAL TAGGAGGGATAL 3110 GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG	AGGCCC MAGGCCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium .intracellulare .paratuberc .tuberculosis .bovis
638 3283 3984 570 3376 3307 1910 3011 5462 3322 677 3322 4023 609	CAAGATCAGGTT CCAAGATCAGGTT CCCGCAGACCAGAC	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC	TTAGA GGGATAL TTAGAGGGATAL TTAGAGGGATAL TTAGAGGGATAL TAGGAGGGATAL TAGGAGGGATAL TAGGAGGGATAL 3110 GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG	AGGCCC MAGGCCC MAGGCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium .intracellulare .paratuberc .tuberculosis .bovis .phlei
3322 677 3322 4023 609 3415 3309 1910	CAAGATCAGGTT CCAGCAGACCACCCGCAGACCACCCGCAGACCACCCCCCC	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC	TTAGASGGATAS TTAGAGGGATAS TTAGAGGGATAS TTAGAGGGATAS TAGGAGGGATAS TAGGAGGGATAS TAGGAGGGATAS TAGGAGGGATAS TAGGAGGGATAS TAGGAGGGATAS 3110 GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGTCAGACCTGG	AGGCCC MAGGCCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium .intracellulare .paratuberc .tuberculosis .bovis .phlei .leprae .gastri
3322 677 3322 4023 609 3415 3305 3050	CAAGATCAGGTT CCAGCAGACCA CCCGCAGACCA	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCACAC ACGGGATTGATA ACGGGATTGATA ACGGGATTGATA ACGGGATTGATA	TTAGASGGATA TTAGAGGGATA TTAGAGGGATA TTAGAGGGATA TTAGAGGGATA TAGGAGGGATA TAGGAGGGATA TAGGAGGGATA 3110 GGCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG	AGGCCC MAGGCCC MAGGCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium .intracellulare .paratuberc .tuberculosis .bovis .phlei .leprae .gastri .kansasii
3322 677 3322 4023 609 3415 3305 3050	CAAGATCAGGTT CCAGCAGACCAGAC	TT-CTCACCOTT TT-CTCACCCTT TT-CTCACCCAC TT-CTCACCACAC ACGGGATTGATA ACGGGATTGATA ACGGGATTGATA ACGGGATTGATA	TTAGASGGATA TTAGAGGGATA TTAGAGGGATA TTAGAGGGATA TTAGAGGGATA TAGGAGGGATA TAGGAGGGATA TAGGAGGGATA 3110 GGCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG GGCCAGACCTGG	AGGCCC MAGGCCC MAGGCC	.intracellulare .paratuberctuberculosis .bovis .phlei .leprae .gastri .kansasii .smegmatis .avium .intracellulare .paratuberc .tuberculosis .bovis .phlei .leprae .gastri .kansasii

Figure 4L

		130	140	150	160
107	GAGTAAC	ACGTGGGC	AATCTGCCC	TGCACTTC-6	GGATAA M.avium
59	GAGTAAC	ACGTGGGC	AATCTGCCC	TGCACTTC-6	GGGATAA M.intracellula
107	GAGTAAC	ACGTGGGC	AATCTACCC	TGCACTTC-6	GGGATAA M.paratuberc.
70	GAGTAAC	ACGTGGGC	AATCTGCCC	recaemic-e	GGGATAA M.scrofulaceum
70	GAGTAAC	ACGTGGGT	danchecce.	7GCACTTC - 0	GGATAA M. tuberculosis
209	GAGTAAC	ארפייפפפיים	CATOTOCO	rGCACTIC - C	GGATAA M. bovis
120	GEGTEE	מרפייפפפיול	בן הדטופטטט. מאייריים כיכים		GGATAA M.leprae
69	CACTANC	ACG1GGGE	MATCIGCCC.	recreated a	GGATAA M.Ieprae GGGATAA M.kansasii
70	CACTAAC	'ACG1GGGC	NATCIGCOC.	rccacacc-e	GGATAA M.Kansasii GGGATAA M.gastri
104	CACTARC	ACGIGGGC		raca careca	GGATAA M.gastri
64		ACG 1GGGTh	ANTO TOCCO	recacarre-e	GGATAA M.gordonae
24	GAGTAAC	ACGTGGGQ	ATCTGCCC:	rgcacttc-e	GGGATAA M.marinum
	•				
					
		450	460	470	480
124	AAACCTC	TTTCACCAT	CGACGAAGG	TCCGGGTTT	TTCTCGG M.avium
376	AAACCTC	TTTCACCAT	CGACGAAG	TCCGGGTTT	TTCTCGG M.intracellule
124	AAACCTC	TTTCACCAT	CGACGAAGG	TCCGGGTTT	TCTAGG M.paratuberc.
387	AAACCTC	TTTCACCAT	CGACGAAGG	CTCACT	TTTGTGG M.scrofulaceum
889	AAACCTC	TTTCACCAI	CGACGAAGG	TCCGGGTTC	TCTCGG M.tuberculosis
528	AAACCTC	TTTCACCAT	CGACGAAGG	TCCGGGTTC	TCTCGG M.bovis
139	ABACCTC	TTTCACCAT	CCACCAACC	TOUCKE AND	TTCTCGG M.leprae
386	AAACCTC	тттсъссъч	CCACCAACC		TCTCGG M.kansasii
387	PARCOTO	アヤヤとなってなり	rcencennee	TCCGGG11C	TCTCGG M.gastri
120	AAACCTC	T T T CACCA!	rcdaceaace	TCCGGG1 IE	TCTCGG M.gastri
81	AAACCIC	**************************************	CCACCAACC		TCTCGG M.gordonae TCTCGG M.marinum
01	AAACCIC	111CACCA	CGACGAAGG	alficeeea.i.i.i	TCTCGG M.marinum
		400	500		
_		490	500	510	520
9					TACGTG M.tuberculosis
8	ATTGACGG	TAGGTGGA	GAAGAAGCA	CCGGCCAAC	TACGTG M.bovis
4	ATTGACGG	TAGGTGGA	GAAGAAGCA	CCGGCCAAC	TACGTG M.avium
6					TACGTG M.intracellula
4	ATTGACGG	TAGGTGGA	GAAGAAGCA	CAC	TACGTG M.paratuberc.
4	GTTGACGG	TAGGTGGA	GAAGAAGCA	CCGGCCAAC	TACGTG M.scrofulaceum
9	ATTGACGG	ТАССТССА	GDAGAAGCA	CCGGCCAAC	TACGTG M.Scrotulaceum
6	ATTGACGG	ТДССТССХ:	CDACDBCCB	CCGGCCDAC	TACGIG M. Leprae
7	ATTCACCC	TACCICCA	CARCARGOR	~~GGCCAAC	TACGTG M.kansasii TACGTG M.gastri
ó	EGENCES	TAGGIGGA	これれなみれなくれ	CCCCCCTTC	TACGTG M.gastri TACGTG M.gordonae
1	PAT GWCGG	TAGGIGGA	プレンジをひたひかれた アファママママ	CCGGCCAAC	TACGTG M.gordonae
_	ALIGACGG	TAGGTGGAI	ЭАНОАНЬСА	CCGGCCAAC	TACGTG M.marinum

Figure 5A

	1130	1140	1150	116	0
1104	TCTCATGTTGCCA	GGGGTAATGC	CGGGGACTCG	TGAGAG	M. avium
1056	TCTCATGTTGCCA	GCGGGTAATGC	CGGGGACTCG	הממממת	M intracellulare
1098	TCTCATGTTGCCA	GCGGGTAATGC	GGGGACTCG	TGAGAG	M. paratubero
1064	TCTCATGTTGCCA	GCGGGTAATGC(CGGGGACTCG	TGAGAG	M. scroful aceum
1069	TCTCATGTTGCCA	GCACGTAATGG	GGGGACTCG'	TGAGAG	M.tuberculosis
1208	TCTCATGTTGCCA	gcacgtaatggt	GGGGACTCG'	TGAGAG	M. bovis
1119	TCTCATGTTGCCA	GCACGTAATGGT	GGGGACTCG'	TGAGAG	M.leprae
1066	TCTCATGTTGCCA	GCGGGTAATGC	GGGGACTCG'	TGAGAG	M. kansasii
1067	TCTCATGTTGCCA	GCGGGTAATGC	CGGGGACTCG'	TGAGAG	M.gastri
1100	TCTCATGTTGCCA	GCGGGTAATGC	CGGGGACTCG	TGAGAG	M. gordonee
1061	TCTCATGTTGCCA	GCACGTAATGG1	GGGGACTCG	rgagag	M.marinum
			•		
		_			
		······································			
	1290	1300	1310	132	0 .
1264	CGAATCCTTTTAA	AGCCGGACTCAG	TTCGGATTE	GGTCT	M.avium
1216	CGAATCCTTTTAA	AGCCGGTCTCAG	TTCGGATTG	GGTCT	M. intracel lulare
1258	CGAATCCTTTTAA	AGCCGGACTCAG	TTCGGATTG	GGTCT	M.paratuberc.
1224	CGAATCCTTTTAA	AGCCGGTCTCAG	TTCGGATOG	GGTCT	M.scrofulaceum
1229	CGAATCCTTA-AA	agccgg ctcag	TTCGGATCG	GGTCT	M.tuberculosis
1368	CGAATCCTTA-AA	AGCCGGMCTCAG	TTCGGATTCG	בכתיירייי	M howis
1279	CGAATCCTTTTAA CGAATCCTTTTAA	agccggictcag	TTCGGATCG	GGTCT	M.leprae
1226	CGAATCCTTTTAA	agccggtctcag	TTCGGATCG	GGTCT	M.kansasii
1227	CGAATCCTTTTAA	agccggtctcag	TTCGGATCG	GGTCT	M.gastri
1260	CGAATCCTTTTAA	agccgghctcag	TTCGGATCG	GGTCT	M.gordonae
1221	CGAATCCTTT AA	agccggiictcag	TTCGGATCGG	GGTCT	M.marinum
		•			
					
	1330	1340	1350	136	0
1304	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.avium
1256	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.intracellulare
1298	GCAACTAGACCOA	ATGAAGTCGGAG	TCGCTAGTA	TCGCA	M.paratuberc.
1264	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.scrofulaceum
1268	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.tuberculosis
1407	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.tuberculosis M.bovis
1319	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.leprae
1266	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.kansasii
1267	GCAACTCGACCCC GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.gastri
1300	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.gordonae
1260	GCAACTCGACCCC	TGAAGTCGGAG	TCGCTAGTA	TCGCA	M.marinum
- •		<u></u>			

Figure 5B

Figure 6

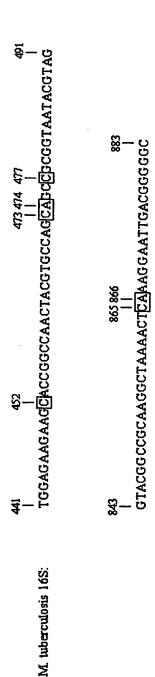


Figure 7

INTERNATIONAL SEARCH REPORT

Inter anal Application No PCT/DK 97/00425

A. CLASSIFICATION OF SUBJECT MATTER IPC 6 C12Q1/68 C07K C07K14/00 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12Q C07K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Category 9 Citation of document, with indication, where appropriate, of the relevant passages Relevant to ctaim No. Y US 5 547 842 A (HOGAN JAMES ET AL) 20 1 - 36August 1996 cited in the application see the whole document Y WO 96 17956 A (GENE POOL INC ; WEININGER 1 - 36SUSAN (US); WEININGER ARTHUR M (US)) 13 June 1996 see the whole document Y WO 95 32305 A (DAKO AS) 30 November 1995 1 - 36see the whole document Α EP 0 572 120 A (GEN PROBE INC) 1 December 1993 cited in the application see the whole document Further documents are listed in the continuation of box C. Patent family members are listed in annex. Special categories of cited documents: "T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance invention "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publicationdate of another involve an inventive step when the document is taken alone document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 20 January 1998 30/01/1998 Name and mailing address of the ISA -Authorized officer European Patent Office, P.B. 5818 Patentlaan 2 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016 Knehr, M

Form PCT/ISA/210 (second sheet) (July 1992)

1

INTERNATIONAL SEARCH REPORT

Inter. Inal Application No PCT/DK 97/00425

C.(Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	PCT/DK 97/00425
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to daim No.
		Present to dain 140.
Ρ,Υ	WO 96 36734 A (ABBOTT LAB) 21 November 1996 see the whole document	1-36

1

INTERNATIONAL SEARCH REPORT

information on patent family members

PCT/DK 97/00425

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
US 5547842 A	20-08-96	US 5541308 A	30-07-96	
		US 5595874 A	21-01-97	
		US 5593841 A	14-01-97	
•		US 5683876 A	04-11-97	
		US 5677127 A	14-10-97	
		US 5677128 A	14-10-97	
		US 5677129 A	14-10-97	
		US 5693468 A	02-12-97	
		US. 5691149 A	25-11-97	
		US 5693469 A	02-12-97	
		US 5679520 A	21-10 - 97	
		US 5674684 A	07-10-97	
		DK 413788 A	23-09-88	
	•	AU 616646 B	07-11-91	
		AU 1041988 A	16-06-88	
•		EP 0272009 A	22-06-88	
		JP 1503356 T	16-11-89	
		KR 9511719 B	09-10-95	
		WO 8803957 A	02-06-88	
WO 9617956 A	13-06-96	AU 4418996 A	26-06-96	
		CA 2206127 A	13-06-96	
		EP 0796344 A	24 - 09-97	
		NO 972611 A	11-08-97	
WO 9532305 A	30-11-95	AU 2522095 A	18-12-95	
·		EP 0760008 A	05-03-97	
EP 0572120 A	01-12-93	AU 4114793 A	29-11-93	
		JP 7506723 T	27-07-95	
		WO 9322330 A	11-11-93	
WO 9636734 A	21-11-96	NONE	عه نیو بیو بیو بیده در جه جه می شد که جه که د	

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

efects in the images include but are not limited to the items checked:
☐ BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
GRAY SCALE DOCUMENTS
LINES OR MARKS ON ORIGINAL DOCUMENT
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
Потигр.

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.